

[Handwritten signature]

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30: _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

FOR OFFICIAL USE ONLY

STAFF USE ONLY	
Searcher: <u>BSmith</u>	Type of Search
Searcher Phone #: <u>308-4477</u>	NA Sequence (#) _____
Searcher Location: <u>CMI-1B17</u>	AA Sequence (#) _____
Date Searcher Picked Up: <u>6/23/00</u>	Structure (#) _____
Date Completed: <u>6/27/00</u>	Bibliographic _____
Searcher Prep & Review Time: _____	Litigation _____
Clerical Prep Time: <u>4</u>	Fulltext _____
Online Time: _____	Patent Family _____
	Other _____
	Vendors and cost where applicable
	STN _____
	Dialog _____
	Questel/Orbit _____
	Dr.Link _____
	Lexis/Nexis _____
	Sequence Systems <u>965302</u>
	WWW/Internet _____
	Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 11:30:45 ; Search time 12463.9 Seconds

(without alignments)
-522.537 Million cell updates/sec

Title: US-09-180-798-1

Perfect score: 6695

Sequence: 1 tctagatgacgaacatcgcc.....attaataattttatggtt 6695

Scoring table:

IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_dal:*
2: gb_da2:*
3: gb_com:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p11:*
10: gb_p12:*
11: gb_p13:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_v1:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_v1:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_dal:*
37: em_da2:*
38: em_hum3:*
39: em_hum4:*
40: gb_p14:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_p13:*
51: gb_p15:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	6691	99.9	6695 5 A67796	A67796 Sequence 1
2	560.2	8.4	1755 5 DC093048	U93048 Daucus caro
3	560.2	8.4	1815 5 A67797	A67797 Sequence 2
4	387.4	5.8	4081 5 A67815	A67815 Sequence 20
5	383.2	5.7	88401 8 F23M19	AC007454 Arabidops
6	378.6	5.7	98471 8 AC012654	AC012654 Arabidops
7	317.6	4.7	96475 7 ATEF17M5	AL035678 Arabidops
8	313.6	4.7	65899 50 AC006436	AC006436 Arabidops
9	250.4	3.7	2089 7 A67827	A67827 Sequence 32
10	217	3.2	86064 5 AB013395	AB013395 Arabidops
11	208.2	3.1	84896 7 AB025639	AB025639 Arabidops
12	191.6	2.9	120787 8 AC002292	AC002292 Genom1c s
13	191.2	2.9	134784 8 F17123	AF16182 Arabidops
14	178.4	2.7	89779 7 AB005234	AB005234 Arabidops
15	176	2.6	110211 50 AC005170	AC005170 Arabidops
16	167.4	2.5	86722 7 AB012245	AB012245 Arabidops
17	146.8	2.2	2278 8 SBRK1	Y14600 Sorghum bic
18	117.8	1.8	94369 8 ATAC009755	AC009755 Arabidops
19	117.8	1.8	103960 8 ATAC011664	AC011664 Arabidops
20	116	1.7	113566 8 AC002330	AC002330 Arabidops
21	110.8	1.7	2640 8 ZMA010166	AJ010166 Zea mays
22	110.2	1.6	93257 8 ATEF84	AL034567 Arabidops
23	110.2	1.6	93353 7 ATEF10M6	AL021811 Arabidops
24	109.6	1.6	65958 7 AB009056	AB009056 Arabidops
25	108.2	1.6	3176 5 E12705	E12705 Arabidops
26	108.2	1.6	3176 8 ATU47029	U47029 Arabidops
27	107.2	1.6	9295 5 E12706	E12706 Arabidops
28	107.2	1.6	9295 7 ATHER	D83257 Arabidops
29	107.2	1.6	77040 50 AC004484	AC004484 Arabidops
30	106.2	1.6	73840 7 AB007644	AB007644 Arabidops
31	105.2	1.6	1952 8 ATU55875	U55875 Arabidops
32	105.2	1.6	94812 44 AC018364	AC018364 Arabidops
33	105.2	1.6	99698 50 AC008262	AC008262 Genomic s
34	103	1.5	81736 7 AB019235	AB019235 Arabidops
35	100.8	1.5	114505 8 F20P5	AC002062 Sequence
36	98.4	1.5	68870 50 AC005312	AC005312 Arabidops
37	98.6	1.5	2808 7 RICEPK1	L27821 Oryza sativ
38	95.6	1.4	2010 8 AF024650	AF024650 Arabidops
39	95.6	1.4	5438 7 AT284202	Z84202 A. thaliana
40	95.6	1.4	59376 50 AC006072	AC006072 Arabidops
41	94.2	1.4	115641 8 ATU95973	U95973 Arabidops
42	94	1.4	107234 8 AC007583	AC007583 Arabidops
43	93.6	1.4	80402 7 AB025636	AB025636 Arabidops
44	93.6	1.4	113567 8 AF058826	AF058826 Arabidops
45	92.8	1.4	80196 8 T32N4	AF162444 Arabidops

ALIGNMENTS

RESULT 1
LOCUS A67796 6695 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 1 from Patent WO9743427.
ACCESSION A67796
VERSION A67796.1 GI:4756622
KEYWORDS
SOURCE
ORGANISM
Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
REFERENCE
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source 1. .6695 Location/Qualifiers
BASE COUNT 1845 a 1182 c 1243 g 2421 t 4 others
ORIGIN
Query Match 99.9%; Score 6691; DB 5; Length 6695;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 acctgggacaaagaagcgcctgatgaactgatttaaccgctgtttattattacac 720
DB 661 acctgggacaaagaagcgcctgatgaactgatttaaccgctgtttattattacac 720
QY 721 ttgatgttcataatgaactaatacttttggttaatttcataatataatacccttcct 780
DB 721 ttgatgttcataatgaactaatacttttggttaatttcataatataatacccttcct 780
QY 781 ctgtgatggcagtgagccctcggttaattagcgaacttaatactatcgtatgatactgta 840
DB 781 ctgtgatggcagtgagccctcggttaattagcgaacttaatactatcgtatgatactgta 840
QY 841 tgccttgacagatgatgtcaacattacacatttgcttgctcaacaataaataaac 900
DB 841 tgccttgacagatgatgtcaacattacacatttgcttgctcaacaataaataaac 900
QY 901 tctattatgtcatcttttttggttaacaaagtaactaattgagtcgtatgtgacagcg 960
DB 901 tctattatgtcatcttttttggttaacaaagtaactaattgagtcgtatgtgacagcg 960
QY 961 atgcattacaacacttaagaaactagcttgcagaatcccaaatgtcctgcagagctggg 1020
DB 961 atgcattacaacacttaagaaactagcttgcagaatcccaaatgtcctgcagagctggg 1020
QY 1021 atccaacccttggaaacccttgacatgggttcaatgtgacatgaacaaatgaacaaatg 1080
DB 1021 atccaacccttggaaacccttgacatgggttcaatgtgacatgaacaaatgaacaaatg 1080
QY 1081 ttaataagatgtagtaacttcccttaataatttttagcaagttacgaataattac 1140
DB 1081 ttaataagatgtagtaacttcccttaataatttttagcaagttacgaataattac 1140
QY 1141 aattggcagaatgtcctttaaataatttttcttaatttctaagcgaagcgaatccta 1200
DB 1141 aattggcagaatgtcctttaaataatttttcttaatttctaagcgaagcgaatccta 1200
QY 1201 tcttaagatcctactgaatttaagacataacattttttaaaaaatcattagagt 1260
DB 1201 tcttaagatcctactgaatttaagacataacattttttaaaaaatcattagagt 1260
QY 1261 gtttttccgacagcgacataatccttttcttgtaattcaagaacaccttcctccg 1320
DB 1261 gtttttccgacagcgacataatccttttcttgtaattcaagaacaccttcctccg 1320
QY 1321 acgataaataataagatlaactccttgaactaatttttaatttcttcttcta 1380
DB 1321 acgataaataataagatlaactccttgaactaatttttaatttcttcttcta 1380
QY 1381 tgtctcttcagagaagttcttcttgcttcttggaagaactcttaataatttt 1440
DB 1381 tgtctcttcagagaagttcttcttgcttcttggaagaactcttaataatttt 1440
QY 1441 tggcaactcatalaataattatataatccatgtagtlaabaagttaaaaaacgtccta 1500
DB 1441 tggcaactcatalaataattatataatccatgtagtlaabaagttaaaaaacgtccta 1500
QY 1501 ttaattccaagatagaggttcatatttatagtttgggcacacacaggttttgaaaaatgt 1560
DB 1501 ttaattccaagatagaggttcatatttatagtttgggcacacacaggttttgaaaaatgt 1560
QY 1561 cagaatcttcttgtagtlaatttacttaacaaactttatgsgtcatgcagatccttg 1620
DB 1561 cagaatcttcttgtagtlaatttacttaacaaactttatgsgtcatgcagatccttg 1620
QY 1621 ggaatgcagaattactgtgcaattgttccctcttgccaggttgaanaattacaact 1680
DB 1621 ggaatgcagaattactgtgcaattgttccctcttgccaggttgaanaattacaact 1680
QY 1681 tgaagaaccatataacttgaatgactttagtttttaacagcacaatgcttcaatact 1740
DB 1681 tgaagaaccatataacttgaatgactttagtttttaacagcacaatgcttcaatact 1740

QY 1741 gctaaagtgtgaaaagtgaacttcttagcttcaagcagtggttcggaataatactatga 1800
|||||
Db 1741 gttaaaagtgtgaaaagttgactttctagctcagcagtggttcggaataatactatga 1800
QY 1801 agcactaaaagcgtcggaacttttttggatattatcttaataatgttaattgttacta 1860
|||||
Db 1801 agcactttaaaagcgtcggaacttttttggatattatcttaataatgttaattgttacta 1860
QY 1861 cttataatgaataactgaatttaactcctcatgatgtgtctcaagtcgaatgtgccctatt 1920
|||||
Db 1861 cttataatgaataactgaatttaactcctcatgatgtgtgtctcaagtcgaatgtgccctatt 1920
QY 1921 agtcaacatnaataaattgunggggttggacaataataaacttcttcttaagttccagaaa 1980
|||||
Db 1921 agtcaacatnaataaattgunggggttggacaataataaacttcttcttaagttccagaaa 1980
QY 1981 gagactatacaactctgtctcagcgcataaagtcacagtggtgtcaagtcgaaggtatcca 2040
|||||
Db 1981 gagactatacaactctgtctcagcgcataaagtcacagtggtgtcaagtcgaaggtatcca 2040
QY 2041 gtttggggagggtttaaagcacttaattacctgtctcttaagcgttcgaagagtta 2100
|||||
Db 2041 gtttggggagggtttaaagcacttaattacctgtctcttaagcgttcgaagagtta 2100
QY 2101 ttaaaagtcgatacatcattcagaagtttaaatagcacttcagtgattatgaatggtaca 2160
|||||
Db 2101 ttaaaagtcgatacatcattcagaagtttaaatagcacttcagtgattatgaatggtaca 2160
QY 2161 tgaagaatcaatcttaattgttctcctatgctgttcaacatgtccttaatactctgtta 2220
|||||
Db 2161 tgaagaatcaatcttaattgttctcctatgctgttcaacatgtccttaatactctgtta 2220
QY 2221 tcttggcactcattaaaatgtgcactgtatataatgctgagaaggtagctcttcaatacca 2280
|||||
Db 2221 tcttggcactcattaaaatgtgcactgtatataatgctgagaaggtagctcttcaatacca 2280
QY 2281 ttcaatgtataccagagaataatcataatltttttaaatcataagttggtccctagagttc 2340
|||||
Db 2281 ttcaatgtataccagagaataatcataatltttttaaatcataagttggtccctagagttc 2340
QY 2341 tctcaagattgtctattatattatcttccacacatttagaacgtgtgttcagatgaataatc 2400
|||||
Db 2341 tctcaagattgtctattatattatcttccacacatttagaacgtgtgttcagatgaataatc 2400
QY 2401 ttggaactccacagagaatcttataagtaaaagtctcttagatcgtatgaaggtttg 2460
|||||
Db 2401 ttggaactccacagagaatcttataagtaaaagtctcttagatcgtatgaaggtttg 2460
QY 2461 tcaatgtgtgtgcccgtccagaaatttaaatcaatcccatgacacatgtttgtatcgtga 2520
|||||
Db 2461 tcaatgtgtgtgcccgtccagaaatttaaatcaatcccatgacacatgtttgtatcgtga 2520
QY 2521 ctactcaactgttaatcgaaaggttaactatgttgaatataatgtttttttttgttctc 2580
|||||
Db 2521 ctactcaactgttaatcgaaaggttaactatgttgaatataatgtttttttttgttctc 2580
QY 2581 catgcttagcgtttaaaaggtctacgtctgacatgtgttttaacatgittatgtttgt 2640
|||||
Db 2581 catgcttagcgtttaaaaggtctacgtctgacatgtgttttaacatgittatgtttgt 2640
QY 2641 actgaagaagttaaaagtcttctctgttttagcaaatgaatatataataaagcgttt 2700
|||||
Db 2641 actgaagaagttaaaagtcttctctgttttagcaaatgaatatataataaagcgttt 2700
QY 2701 aacttctcgttggaaggtgttcttaacttttatataataataataataataagactctc 2760
|||||
Db 2701 aacttctcgttggaaggtgttcttaacttttatataataataataataataagactctc 2760
QY 2761 gctggcaatatatcttcgaactcagatcagatcagaactgtatattaggttcagatg 2820
|||||
Db 2761 gctggcaatatatcttcgaactcagatcagatcagaactgtatattaggttcagatg 2820
QY 2821 agtggcgttagtagaacaacttaagcaagaacttaataatcaatgaaggttcaaacctttaact 2880
|||||

Db 2821 agggcgttagtagaacaacttaagcaagaacttaataatcagatgaggtttcaaccttttaact 2880
|||||
QY 2881 ttcttttagatttttttcaagttatggaanaatgttaactaatgtgcgtgttcttctcc 2940
|||||
Db 2881 ttcttttagatttttttcaagttatggaanaatgttaactaatgtgcgtgttcttctcc 2940
QY 2941 atnaacttccataaagtcggttcttctgaagtttcaatgtgaagtcgttcgaagttat 3000
|||||
Db 2941 atnaacttccataaagtcggttcttctgaagtttcaatgtgaagtcgttcgaagttat 3000
QY 3001 atnaagcgttcttcaataatcaataatgltctcaactttgataagcgttactatbat 3060
|||||
Db 3001 atnaagcgttcttcaataatcaataatgltctcaactttgataagcgttactatbat 3060
QY 3061 tgcaccttgcaacttaactgtatcttctcctaatgtcatcttgatgtcaatgtcaactaaccga 3120
|||||
Db 3061 tgcaccttgcaacttaactgtatcttctcctaatgtcatcttgatgtcaatgtcaactaaccga 3120
QY 3121 gttgaacatggtttatcaatgcttcttggaggtgaacaatgtagcttcaactcgtctctga 3180
|||||
Db 3121 gttgaacatggtttatcaatgcttcttggaggtgaacaatgtagcttcaactcgtctctga 3180
QY 3181 tabaagtttaaggtctgcaactcctccactagccttcgttgtttatcaacagttaacaca 3240
|||||
Db 3181 tabaagtttaaggtctgcaactcctccactagccttcgttgtttatcaacagttaacaca 3240
QY 3241 cctactaagcagcgttccactcctcctccttcttctgcgcaaatgtagaaggttcttctgca 3300
|||||
Db 3241 cctactaagcagcgttccactcctcctcctccttcttctgcgcaaatgtagaaggttcttctgca 3300
QY 3301 taatagtgatcatcatcttaagaataatgtagaataatcatatcgttgttatgtgttctac 3360
|||||
Db 3301 taatagtgatcatcatcttaagaataatgtagaataatcatatcgttgttatgtgttctac 3360
QY 3361 ttggaattaaatgaggtctgcaactgttgcgttgtttatcgtgtgttcaagcgttgcaca 3420
|||||
Db 3361 ttggaattaaatgaggtctgcaactgttgcgttgtttatcgtgtgttcaagcgttgcaca 3420
QY 3421 ggaagcagattgtagaaggtggtggtcccaagcgccttaacttgaagggcggtgcgac 3480
|||||
Db 3421 ggaagcagattgtagaaggtggtggtcccaagcgccttaacttgaagggcggtgcgac 3480
QY 3481 gttgtgattctgggtgtctatcttagaagacatctatctatactatgattttatga 3540
|||||
Db 3481 gttgtgattctgggtgtctatcttagaagacatctatctatactatgattttatga 3540
QY 3541 atataatcaactactatatttcttctgttaatgtttatagaaatcccatcgttaac 3600
|||||
Db 3541 atataatcaactactatatttcttctgttaatgtttatagaaatcccatcgttaac 3600
QY 3601 ttgacaataacatcgtgaatatltgaacctaatgaatgaatgtgtcaggttttaattca 3660
|||||
Db 3601 ttgacaataacatcgtgaatatltgaacctaatgaatgaatgtgtcaggttttaattca 3660
QY 3661 aactcaattcaatttcaataaataatcttataatgaatcgttaagatgaatatat 3720
|||||
Db 3661 aactcaattcaatttcaataaataatcttataatgaatcgttaagatgaatatat 3720
QY 3721 taaatataatgattgtgtccatataatagcgtgaatgtctaaatagaccccaagcagctg 3780
|||||
Db 3721 taaatataatgattgtgtccatataatagcgtgaatgtctaaatagaccccaagcagctg 3780
QY 3781 ctctacgtccagagcgtccagagcaggttcaactgcatgtcttaactctgaacaataatgggttc 3840
|||||
Db 3781 ctctacgtccagagcgtccagagcaggttcaactgcatgtcttaactctgaacaataatgggttc 3840
QY 3841 gtagacattgttggagatccctcaactgagatgtctcttggcagaccctcgtttcaat 3900
|||||
Db 3841 gtagacattgttggagatccctcaactgagatgtctcttggcagaccctcgtttcaat 3900
QY 3901 tgaattcatatgtagtattactagtttataaataatcttcaatttgcaataatttaac 3960
|||||

Db 3901 TGATTTCATGATGATAGATTACTAGTTTATATAATATCTTATGCAATAATTTAAC 3960
QY 3961 tggagtttaacaatgacagggagctttacagcaataaataaagggaccaatttcctgtg 4020
Db 3961 TGGAGTTTAAACATGACAGGGAGCTTTACACCAATTAACATAGGGACCAATTCCTAGTG 4020
QY 4021 atcttgggaatctgcaaatgttggtagcttggacctataatgaatcctctctgagc 4080
Db 4021 ATCTTGGGAATCTGCATAATTTGGTAGCTTGGACCTATACATATATAGCTTCTCTGAC 4080
QY 4081 ctataccggaacatlaagaaagcttacaaggtlaagaattctgtatgactacaactct 4140
Db 4081 CATATCCGGACACATTAGAAAGCTTACAAAGCTTAAGATCTGTATAGCTACAAACTT 4140
QY 4141 cactagttttaactaatgcaattgtattatcctttaagtgattgattatcacaaa 4200
Db 4141 CACTGTTTTTAACCTTAATGCAATTTGATATCTTCAATGATGATTATACACAAA 4200
QY 4201 ttaactggaaggcgtctcaacaacaacactgcctctgtgccaattccaatgtcaactgact 4260
Db 4201 TTACTGGATAGGCGTCTCAACACAACTGCCTCTGTGCTCAATTCCAATGCTCACTGACT 4260
QY 4261 aatatlaacaactctcaagctcgtlaaglatccagaccttccagatagtttgtgtg 4320
Db 4261 AATATTACAACCTCTCAAGTCTGTAAAGTATCCGACCTTCCAGATAGTTTGTGTG 4320
QY 4321 tggatgtttcaatttaactaataatgtaatgtaacagagatttaacaacatgctat 4380
Db 4321 TGGATGTTTCAATTTAAATTAATTAATGATCATCAGGGAATTTTCAAAACATGGGCTAT 4380
QY 4381 cagagaccagtlacaggaataagtgctacttctctgttacaacatacaggttaagtct 4440
Db 4381 CAGGACACAGTACGGATATAGGCTATTTCTTGTATACACATACGATTATATGTA 4440
QY 4441 gtaataccttaataattatgcttactactactactcgtgaaggtcgtatatttctt 4500
Db 4441 GTAAATATCTTAAATTAATGTTTACTTACTACGCAAAAGCTATGATTAATTTT 4500
QY 4501 tctcctcaataataacttgcagtttggcaaatatttaattatgtagctg 4560
Db 4501 TCTCCTTCAATTAATTAATCCTTTCGCAAGTTTGGCAATTAATTAATGAGCACTG 4560
QY 4561 taacttgggaagccctgcctctggatctccccaatttctccaaccaactccgttccac 4620
Db 4561 TAACTGGGAGGCCCTGCCTGTGATCTCCCACTTTCACACACCTCGTCAATCCAC 4620
QY 4621 catcaacagtlacagctccaggtgattagtttataataatccgtaattatctta 4680
Db 4621 CATCAACAGTACAGCCTCCAGGTGATTAATTAATTAATTAATTAATTAATTAATTA 4680
QY 4681 tgactgttaaaaaattggtttaattccacaggttgcgaataaagtatttccctctctc 4740
Db 4681 TGAATGTAAAAATTTGGTGTATTTCACCAAGTGGCAATTAAGTATTTCTTCTTCTC 4740
QY 4741 ttctattatlaagaagcaaatggtcccaacttggagctatctgcggggagtagctgc 4800
Db 4741 TTCTTATTAATTAATGAAGCAAAATGTCCTCCTGAGCTATGCTGGGGAGTAGTGC 4800
QY 4801 tgggtgctgttactggttgcctgcacccgcaatgycattgcaatggtgggagaagaaa 4860
Db 4801 TGGTGTGCTTAACTGTGTGCTGCACCTGCATGGCATTGTGATGGTGGCGAGAAAGAA 4860
QY 4861 accgcgagaacattccttggatgctgcaggttagtcctgttaaatgatactattgaagc 4920
Db 4861 ACCGCGAAGAACTTCTTTGATGTGCCAGTTAGTCCGTAAATTAATTAATTAATTAAGC 4920
QY 4921 gcttactgctgtagacttgttttcaactgltcaatagtttaactcagctgaagagacc 4980
Db 4921 GCTTACTGCTGTGGAGCTTGTTCACGTCATAGTTAATTAATTAATTAATTAATTAAGC 4980
QY 4981 agaaagtcacacttgttcaactgaagaggtttctctggaggaattgcaagtcgcaagga 5040
Db 4981 AGAAAGTCACCTTGTGCAACTGAAGAGTTTCTGTGCGAGAAATGCAAGTCGCAAGCGA 5040

QY 5041 tacttttagtaccatccttggagaaggtgatttggtaaggtgtataaaggagcccttgc 5100
Db 5041 TACTTTTAACTACATCTTGGAAAGGTGGATTTGGTATGATTAAGGAGCCCTTGC 5100
QY 5101 tgatgctcaacttgaagcaattaaaggctlaaagaagaagcaacaggttgcagact 5160
Db 5101 TGATGGCTCACTGTATGACGATTAAAGGCTTAAGAAAGAAAGAACACAGAGTGGCGAGCT 5160
QY 5161 gcaagttcaacaagaagtggaatgattagcaatgctgtgcacggaatcctctgcct 5220
Db 5161 GCAGTTTCAAAACAGAGTGAATGATTAAGCATGCTGTGCATGGAATCTTCTGCGCT 5220
QY 5221 acgtgttcttgcatagaacccctcagcagccttcttataatcaatcaggttaatg 5280
Db 5221 ACGTGTTCTGCATGACACCTACCGAGCGGCTTCTGTATATCCATCATAGGCTAAATGG 5280
QY 5281 aagttgctcatgtttaagaggtatcagttacaattacaataacttgcagaagt 5340
Db 5281 AAGTGTGGCTCATGTATTAGAGGTATCTCAGTATCAATTAACATCTCCAGAGTT 5340
QY 5341 tgttgattaaaaatgaaatataactccctacataatgtaaggtgtaattcttgag 5400
Db 5341 TGTGTGATTAATAAATGAATATTAATCTCCATCACTATGTAAAGGTATTAATTTCTGAG 5400
QY 5401 cagatctatcccatgcaagatacagaattattgttttcttctgtaattgatacgg 5460
Db 5401 CAGATTTATTTCCCATTCGACAGATACCAATTAATTAATTTCTTCTGTAATTAATACG 5460
QY 5461 gtataattcttcttctgtlaattgttataatgaagatttcgaagcttaataatga 5520
Db 5461 GTATATTTCTTCTTCTGTATTTGTTATATGCAAGATTTCGAGTCAATTAATTAATCA 5520
QY 5521 actggaatgctatgattatctcgaattgaattcctgtcctatgycgcaaatatata 5580
Db 5521 ACTGAGTCAATGATTAATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATGA 5580
QY 5581 ttcaacttgaatcatctataataatactatgtaagtcagctgttgaacttcaatla 5640
Db 5581 TTCAACTTGAATCATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5640
QY 5641 attagcttcaataatcagaatctgcctagtagcttaccgacatactcaaaccttcc 5700
Db 5641 ATTAGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5700
QY 5701 ttaagccctgtatataatcgtcccaacttacttattgaatttgcctctcgaatt 5760
Db 5701 TTAAGCCCTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5760
QY 5761 ttgactgttacaattgtgattctgttttcaatcaaatgtagcgtccagcatcagaacc 5820
Db 5761 TTGATGTGATCATGTGATGATCTTGTTCATCAATTAAGAGCTCAAGCATCAAGAAC 5820
QY 5821 tccccctgattggccaactaggagaagattgacatagatcttccagggtcctactaa 5880
Db 5821 TCCCCCTGATTGGCCAACTAGAGGAGAGATTGCACTAGAGATCTTGAAGGCGCTATCAA 5880
QY 5881 attgcatgaccatttggatcccaagaattatcactgcgattgaaagtgcaaatatact 5940
Db 5881 ATTGCAATGACCATGTGATCCCAAGATTATTCATGCGATGTAAGTGTGAATTAATTT 5940
QY 5941 attgagcagaagaatttgaagcgttgttgaatgatttggtaggttgaagcctatgagta 6000
Db 5941 ATTGAGCAGAAATTTGAAGCGTGTGTAGGTATTTTGGGTTAGCTAGGCTCATAGATTA 6000
QY 6001 caagatacccaatgtaagcgtcgttaagggtaaccatttgggcaacatgctcccgatla 6060
Db 6001 CAAGGATACCAATGTATACAGTGTGTAAGGGGTACCAATTTGGGACATAGCTCCGAGTA 6060
QY 6061 cctctgactggaaggtatctcagagaagcagatgtcttggttatggataatgctcct 6120
Db 6061 CCTCTGACTGAGAAAGTCAATCAGAGAACGATGTTTGGTTATGGGATTAATGCTCCT 6120

QY 6121 agagcttacttgcagagagagcttctgacttgcgccttgcagacgatgatgagt 6180
 |||||
 Db 6121 AGAGCTTACTTACTGAGAGAGGGCTTTGATCTTCTCCCTGCGAAGATGATGATG 6180
 QY 6181 tatcttgatgatgagtgatgctccgggtgtctcttggttaattatctacatata 6240
 |||||
 Db 6181 TATCTTGTGGATTTGGTATGCTGCTCCGGGTGCTCTTGTTATTTATTTTCACTATTA 6240
 QY 6241 gtgtctactacttctgtgtgtcccttctgtttatttccgcgcgtattgattctagt 6300
 |||||
 Db 6241 GTGCTTACTTCTTGTGTGTGGCCCTTGTGTTTATTTCTCCGCTGTATTTGATTTCTAGT 6300
 QY 6301 catgttatgatatttgaccgtcttgcaatgctctttaggttaaaagcctttgaaag 6360
 |||||
 Db 6301 CATGTTATGCAATTTGACCTGCTTGTGCAATGCTTTAGGTTAAAGCCTTTGAAAGAG 6360
 QY 6361 aaaaagttgagatgctgcgtctgcacccctgcacctgcagaaacatcatcatcagaagt 6420
 |||||
 Db 6361 AAAAAGTTGGAGATGCTGTGCTGATCTGACCTGCAGAACATTTACATTGACACAGAGTT 6420
 QY 6421 gagcagcttatctcaagtacatctctgtaccccaaggtctgcacatgagcgccttaag 6480
 |||||
 Db 6421 GAGCAGCTTATTCAGTAGACATTACTCTGTACCCAGAGGTTCCCAATGAGCGGCTTAAG 6480
 QY 6481 atgtcagaagtagtccgaatgcttgaagtgatgctgcgtctgcagaaagtggagcagtg 6540
 |||||
 Db 6481 ATGTCAAGAGGTATGCCAATAGCTTGAAGGTGATGGCTTGCANAAAAAGTGGAGCGATGG 6540
 QY 6541 caaaaagttgaatgcatcatcacaagacgtagaattagctccacatcgaactctgaaag 6600
 |||||
 Db 6541 CAAAAGTGTGAAGTCATCCATCAAGAGGTGAATATGCTCCACATGCAATCTCTGATATG 6600
 QY 6601 atcttagactgcagaagtaactgcacgtcttcttgaattatctgtgtcagaagtaaaagca 6660
 |||||
 Db 6601 ATCTTAGACTGCAAGATTAATCTGCATGCTTTGATTAATTAATCTGTCACAGATTAACAGCA 6660
 QY 6661 tataaatgtaatgaattaatatttttatggtt 6695
 |||||
 Db 6661 TATAAATGTATGATAATATATATTTTATGTT 6695

RESULT 2
 LOCUS DCU93048 1755 bp mRNA PLN 28-JUN-1997
 DEFINITION Daucus carota somatic embryogenesis receptor-like kinase mRNA,
 complete cds.
 ACCESSION U93048
 VERSION U93048.1 GI:2224910
 KEYWORDS
 SOURCE .
 ORGANISM Daucus carota
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 REFERENCE
 AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
 TITLE A leucine-rich repeat containing receptor-like kinase marks somatic
 plant cells competent to form embryos
 JOURNAL Development 124 (10), 2049-2062 (1997)
 MEDLINE 97313247
 REFERENCE
 AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
 TITLE Direct Submission
 JOURNAL Submitted (12-MAR-1997) Molecular Biology, Agricultural University
 of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
 FEATURES
 source
 1..1755
 /organism="Daucus carota"
 /db_xref="taxon:4039"
 34..1695
 /note="SERK"
 /codon_start=1
 /product="somatic embryogenesis receptor-like kinase"

/protein_id="AAB61708.1"
 /db_xref="gi:2224911"
 /translation="MNRNSIILNLYMOFTADLYDKYGLVLTLEYSNISGPIPSDGL
 /NLAIVSLDLYMNSFSGIPDITLGLKILRLPFLRNNSLSGPIPNLSLNTITLYOVL
 SNMRSGVPNGSFSLEFPTSPANNLNLGPGVTRGPCSPSPSPSPSPSPSPSPSP
 PCGNPTGAIAGGYAAGALLFAPAPAAFWMMRRRRREFRPVPAEDPEVHLGOK
 RFLRELVATDIPSTILGRGGRKVTGKLAGSVAVARKKEETPEGELOFOTRY
 EMIISVAHRNLLRLNGFCMTPTERLLVYPYMWAGSVASCLREKQSPSEPLDPTKRI
 ALGSARGLSYLHDHCDPRILIRVDKANIILDEFEAVVDGFLARLMDYKDTHTTA
 VSGTLGYLAPYLSLSTGKSEKTEVFGYGLIMLELITGQRAFDLRLANDDDVALLDWY
 SLKLERKLEMLVDPDLNNYIDIEVEQLIOVALLTCOGSPMRKMSVVRMLEGDG
 LAEKWDEKXEVFIHDELAHPHTSMIISDNLHAFELSGPR"
 BASE COUNT 506 a 347 c 407 g 495 t
 ORIGIN

Query Match 8.4%; Score 560.2; DB 8; Length 1755;
 Best Local Similarity 81.9%; Pred. No. 1.5e-86;
 Matches 725; Conservative 0; Mismatches 18; Indels 144; Gaps 2;

QY 5800 agagcgctcagccatcagaaaccccccgtatgtgccaaactgagagagatgactagg 5859
 |||||
 Db 984 AGAGCGTCAACCATCGAAGCCCTCCCTTGATGGCAACTGGAAGAGATTGGACTAGG 1043
 QY 5860 atctctaaggggactatcctaattgcatgaccattgtatcccaagatcatcatcgca 5919
 |||||
 Db 1044 ATCTGCTAGGGGCTTCTTATTTCATGACATGCTGATCCCAAGATTATCATCTGTA 1103
 QY 5920 tgaataagctgcacaatatatttggacgaagaatttgaagcgtgttgaagtattgg 5979
 |||||
 Db 1104 TGTAAAGCTGCAAAATATATTTGACCAAAATTTGAGCTGTGTGAGCTGATTTGG 1163
 QY 5980 gtttagcagctcattgattgattacaagagatcccaatttcagactgctgtgaaggatacat 6039
 |||||
 Db 1164 GTTAGCTAGGCTCATGATGATTAACAAGATCCCATGTTACACTGCTGTAAAGGTTACTT 1223
 QY 6040 tgggacatagctcccgagtlaccctcgactgcagtggaagtcatacagaagacgatgtctt 6099
 |||||
 Db 1224 GGGCTACATGCTCCCGAGTACCTCCGACTGGAAGATCATCAGAGAGAGCGATGCTT 1283
 QY 6100 tggttatggagatagtctccttagagctcattacttagaagagaggttttgaattcttcgcg 6159
 |||||
 Db 1284 TGGTTATGGGATTTATGCTCTTAGAGCTCATTTCTGAGACAGAAAGCTTTGATCTTGCTCG 1343
 QY 6160 ccttgcgaagatgcatgattgttctgttgattgggtatgtgccgggtgttccttt 6219
 |||||
 Db 1344 CCTGCGAAGATGATGATGTTATGTTGTGATG----- 1380
 QY 6220 ggttaattattacatatagtgcttaactacttgggtggcccttggtttatttcc 6279
 |||||
 Db 1380 ----- 1380
 QY 6280 tgcctgatttgattcttagtcatgttatgcatattgacatgcttgcgaatgctttag 6339
 |||||
 Db 1380 -----g 1380
 QY 6340 gtttaagccttttgaagaagaaagttggagatgctgctgcatcctgcagac 6399
 |||||
 Db 1381 GTTAAAGCTTTTGAAGAGAAAGGTTGGAGATCGTGCTGATCCCTGAGAGAAC 1440
 QY 6400 aattacatgacacaagaatttgcagcttattcaagtacatgacatcctgttaccagaggt 6459
 |||||
 Db 1441 AATTAATTAACACAGAGAGTGTGACGCTTATTCAGATGACTTACTGTATCCAGAGGT 1500
 QY 6460 tcgcgaattgagcggcctcaagatgctcagaggtgctcgaatgcttgaagtgatgacctt 6519
 |||||
 Db 1501 TCGCAATGAGCGGCTTAAGATGTGAGATGTCGCAATGCTTGAAGGTGATGGCTTT 1560
 QY 6520 gcaagaaagtggagcagtgycgcaaaagtgtgaatcatcatcagaagatgataact 6579
 |||||
 Db 1561 GCAGAAAAGTGGGACGATGGCAAAAAGTAAAGATCATCAAGAGATGATTAAGTACT 1620
 QY 6580 ccacatcgaaactctgaatgagatccagactgcagacagataactgcatgcttgaatta 6639

Db	1621	CCACATCGAAGCTTCGAAATGCAATCTCTAAGCTCGACAGATTAACCTTGATGCTTTGAATTA	1680
Qy	6640	tctgtcccaagaataacagcatataaatgt-aatgaataatattttttatgtct	6695
Db	1681	TCGTGCTCAAGATTAACACATATTAATGTAATGAATTAATATTATTATTTATGCTT	1737
RESULT	3		
A67797	LOCUS	A67797 1815 bp DNA	PAT 05-MAY-1999
A67797	DEFINITION	Sequence 2 from Patent WO9743427.	
A67797	VERSION	A67797	
A67797.1	KEYWORDS	GI:4756623	
ORGANISM	SOURCE	carrot.	
ORGANISM	ORGANISM	Daucus carota	
REFERENCE	REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiales; Daucus.	
REFERENCE	REFERENCE	1 (bases 1 to 1815)	
REFERENCE	REFERENCE	De V.S., Schmidt E.D., Van, H.G. and Hecht, V.F.	
REFERENCE	REFERENCE	PRODUCTION OF APOMITIC SEED	
REFERENCE	REFERENCE	PATENT: WO 9743427-A 20-NOV-1997;	
REFERENCE	REFERENCE	CIBA GEIGY AG (CH)	
FEATURES	FEATURES	Location/Qualifiers	
source	source	1..1815	
source	source	/organism="Daucus carota"	
source	source	/db_xref="taxon:4039"	
source	source	94..1755	
source	source	/note="unnamed protein product"	
source	source	/codon_start=1	
source	source	/protein_id="CAB42241.1"	
source	source	/db_xref="gi:4756624"	
source	source	/translation="NNRNSINILNYMOFTDAYIDKYGVLMTLELYSNNSISGPISPDGLQ	
source	source	NITNTVSDIYVMSRSGPIPTDGLKLRLEPFLNNNSISGPIPSINITTILOVDI	
source	source	NNRNSGVPDNGSPSLPTIPISFANNLITCGPTGRCSPSPSPPTPISTVOC	
source	source	PCONGTGALMGVAAGALFLPAPAFPMWRKRPREHFVDVREDPEVHLQGL	
source	source	RSLSRLAVLDTFSTLILGRGFKVYKGLAGSIVAVKRLKEETPEGELQFTE	
source	source	EMISNAVHRNLRLRGFCMPTERLLVYPWAGSVASCLRQSPSEPLDPTRR	
source	source	AGSARGSLYHDHCDPKIIRDYKANIILIDEEFAVVDGFLANDYKDTHTT	
source	source	VGTSLGYIAPETLSGKSEKTEDEYGINLDELITGORAFDLARLADDVMLDWM	
source	source	KLAKERKLEMLVDDLENNYIDEEVQLIYVALICTGSPMERKSESVYRMLEDG	
source	source	LAEKDEMRKVEYIHQDEYELAPHRTEMLIDSDNLHAFELSPR	
BASE COUNT	531 a 354 c 415 g 515 t		
ORIGIN			
Query Match	8.4%; Score 560.2; DB 5; Length 1815;		
Best Local Similarity	81.9%; Pred. No. 1.5e-86;		
Matches 733; Conservative	0; Mismatches 18; Indels 144; Gaps 2		
Qy	5800	agaagctcaacccatcagaacctccccctgattgagccaactagagagagattgactagg	5859
Db	1044	AGAGCGTCAGCAGCATCAGAACCTCCCTGATTGGCCCACTAGSAGAGAGATTGCATAG	1103
Qy	5860	attcttcaggggcctactcaaatgatgatgacatgtagatcccaagattatccatcgga	5919
Db	1104	ATCTGCTAGGGGGGCTTTCTTATTATTGATGATGACCATGTGTGATCCCAAGATTATCATCTGA	1163
Qy	5920	tgtaaaagctcacaatatattattgagcaagaattttaggctgtgtgaggttgaatttgg	5979
Db	1164	TGTAAAGCTCGCAATATATTATTATGGACGAGATTTGAGGCTGTGTGAGGTGATTTGG	1223
Qy	5980	gttagctagctcatgattacaagaataccatgcttacagctgctgtaagggttaccat	6039
Db	1224	GTTACTAGCTCTCATGATTACAGAGATACCCATGTTACACTGCTGTGAAGGGGTACTTT	1283
Qy	6040	tggagcaatgctccgaagactcctccagctgtagaagatataagaagaagacagctgctt	6099
Db	1284	GGGCTACATTAATCTCCCAAGTAACTCTTGACTGAGAAAGTATATAGGAAGACCGATGCTTT	1343
Qy	6100	tgttatatgggtataatgctccttagagctcatcttacttgtagacagaaggcttttgaattctgctgcg	6159

	Db	1344	TGGTTATGAGATTATACCTTCTTAGACCTCATTAAGCAGACAGACCTTTGTATCTGTCG	1403
Oy	6160	ccttgcgaaacgatgatgaatgttatcttgtttgattggtaagtatgctccggagtgctcctt	6219	
Db	1404	CCTTGGGAACGATGATGATGATTTATCTTGCGATG-----	1440	
Oy	6220	ggttaattattcacataatagtctactactctgttgtgccccttgtttatttc	6279	
Db	1440	-----	1440	
Oy	6280	tgcctgtatttgattccttagctatgcatattgaacccttgcaatgcttttag	6339	
Db	1440	-----G	1440	
Oy	6340	gttaaaagccttttgaagaagaaaaagtgtgagatgctgcgcgacctgcacagAAC	6399	
Db	1441	GTTAAAGCCTTTTGGAAAAGAAGAAAAAGTTGGAGATGCTGCTGCATCCGACCTGGAGAAC	1500	
Oy	6400	aattacattgcacacaagaattgaagcagcttattcaagtaagcattactctgtaccagggt	6459	
Db	1501	AATTACTTTACACAGAAAGTTGAGCAGCTTTATCAAGTACCATTAACCTGTACCAAGGT	1560	
Oy	6460	tccgccaatgagagcgccctaagaatgccaagaagagcgcgaagccttgaagtgatgagcctt	6519	
Db	1561	TCGCCAATGAGAGCGCCTTAAGATGTCAAGAGGTATCCAGAAAGCTTGAAGAGTGAGCGCTT	1620	
Oy	6520	gcagaaaagcggagcagcagcgtgcacaaaagtttgaagtcacatccatcaagacgtlaagtagct	6579	
Db	1621	GCAGAAAGTGGGACGACGATGGCGCAAAGATGAGAAAGTCAATCCATCAAGACGTGAATTAGCT	1680	
Oy	6580	ccacatctgaactctctgaatggatcctagaactgcacagataactgtgcagctttggaatta	6639	
Db	1681	CCACATCGAACTTCTGAATGGAGCTCAAGACTCGACACGATGAATCTTGCAAGCTTTGAATA	1740	
Oy	6640	tcctgtccaaagtaaacagcataaaagt-aatgaaattaatttttatggt	6695	
Db	1741	TCCTGCTCAAGATTAACGACATTAATAAGTGAATGAATTAATTTTTATGCTT	1797	
RESULT	4			
LOCUS	A67815	4081 bp	DNA	PAT
DEFINITION	Sequence 20 from Patent WO9743427.			05-MAY-1999
ACCESSION	A67815			
VERSION	A67815.1	GI:4756638		
KEYWORDS				
SOURCE				
ORGANISM				
	thalie cross. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi- s. (bases 1 to 4081) De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F. AUTHORS TITLE PRODUCTION OF APOMICRIFIC SEED JOURNAL Patent: WO 9743427-A 20-NOV-1997;			
FEATURES				
Source				
	Location/Qualifiers			
	I..4081			
	/organism="Arabidopsis thaliana"			
	/db_xref="taxon:3702"			
	/clone="ARABIDOPSIS SERK GENE"			
exon	1280..1367			
exon	1796..1928			
exon	2014..2085			
exon	2203..2346			
exon	2450..2521			
exon	2617..2688			
exon	2772..2884			
exon	3015..3146			
exon	3305..3646			
exon	3760..4081			

TITLE Direct Submission
JOURNAL Submitted (04-JUN-1999) Plant Gene Expression Center, 800 Buchanan
street, Albany, CA 94710, USA
REFERENCE 4 (bases 1 to 88401)
THEORETICAL
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1999) Plant Gene Expression Center, 800 Buchanan
St., Albany, CA 94710, USA
COMMENT On Jun 17, 1999 this sequence version replaced gi:4996902.
The sequence of BAC F23M19 from Arabidopsis thaliana chromosome 1.
FEATURES
source location/Qualifiers
1. .88401
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F23M19"
complement(796. .3891)
/gene="F23M19.1"
complement(join(796. .1116,1205. .1271,1349. .1432,
1506. .1636,1794. .1850,1999. .2047,2166. .2251,2468. .2545,
2771. .2893,2990. .3079,3154. .3254,3332. .3393,3467. .3556,
3638. .3891))
/note="F23M19.1"
/note="Contains a PF100561 alpha/beta hydrolase fold
domain."
/codon_start=1
/evidence=not_experimental
/protein_id="A039602.1"
/db_xref="GI:5091614"
/translation="MEEDLILSPSPYDLFQALTLPIRHYLALFVLCITFEYNY
EFHGLPAIATRYFRCAVNLINPSPSLHYGVSRQILHGRVATPMLASPIHICFL
NHEGLPVTYRQLFLSDGFIALDMLTNSDVIDGSLNHESEITKEDTPIVAVP
GLTSSSSATIKLADYDPAKTGNNYISNHRGLGCVATSCDFYNAKGTEDTPIVAVP
LQHYTPAPLPAIGTISIGANVAVKYLGESEKTPLEGAVALSCPMDLIGRCITRL
KQKLYDALTIGLQTAQLHPQLPLAWEGIKKRSRIIDDMNATLGVKFEFVLP
YKRSSTOYGVAVAPLCLISALDPLCTKRAIPWDECRANKIVLATYHGHLEF
FEGLTGSSLMWVATINFLGLVLSGSPYMIQKIVDKSSGSGQEPSINQPYNTLAE
DGLVAAYKEDQDKTKTKLQGRKRPEDVTKKSFELCKHOTKOSVWMLGYIGMTSP
LLGLMNYIFRKKORPTTSTS"
complement(5693. .6732)
/gene="F23M19.2"
complement(join(5693. .6108,6233. .6428,6499. .6732))
/note="Similar to gb|X98322 peroxidase (pxr10) from
Arabidopsis thaliana."
/codon_start=1
/evidence=not_experimental
/protein_id="A039603.1"
/db_xref="GI:5091615"
/translation="MRAITASFTEFCFVLVPSVFGKIRGFTDSSCPLEASTVSRVAVK
HHSLNQVTALAMQFHCFFVVRKLLLCVYVGGDASLLDSTPEPSEKSAEANY
SVRGFEIIDEKKELEIVCPKVSQADIVYLAIKDSIALAGLANTNNVALIGGHSY
GVACSLFQDRKDPKMSDLKAKLQNTCRGPDPSVLDQATPLEVNOYAIKQSKQ
RGILRIDQNLGLDSTSRISVSNALMNTLFGERFAAMQIMGEIVVLGNSEIETNC
IVVSH"
complement(10781. .14114)
/gene="F23M19.3"
complement(join(10781. .11318,11416. .11460,11670. .11746,
11875. .11991,12207. .12310,12443. .12512,12614. .12717,
12848. .12941,13174. .13213,13394. .13719,13800. .14114))
/gene="F23M19.3"
/codon_start=1
/evidence=not_experimental
/product="F23M19.3"
/protein_id="A039604.1"
/db_xref="GI:5091616"
/translation="MGGLSRSSSVNNAPGQTPAHVNGHLLNNNASDLNSHGESGLK
DDPSVTEVDNKNKTESPSPVYSSGPNIEGIPRLSRVLSOKRSRKSROAA
VAKYSEVSLIGRAGTGIGKAVDIVDTGSSMTNLNLSGFSATYAKGNTSILSP
EVANTIVKGAULHMSLSKDSITLHKVEYLPSESVQULISKDDELIRIARAADKQHD
FVTLVTVQRDLIHEHLALDPRFEDDYORKIOEENSPVTAQGVGDTLALIRTELKSKOK
HVRNLKRKSLMSRLIEEMERLVVYVHFLLETHEFAGADPDKPAMPDPIHKKRLGS

gene
CDS
/gene="F23M19.4"
join(23307. .23355,23431. .23627,23656. .23751,23850. .23906,
24002. .24145,24225. .24309,24401. .24491,24571. .24735,
24842. .24955,25045. .25117,25210. .25348,25453. .25865,
25953. .26140,26231. .26279)
/gene="F23M19.4"
/note="Similar to gb|AF082176 auxin response factor 9 from
Arabidopsis thaliana."
/codon_start=1
/evidence=not_experimental
/protein_id="A039615.1"
/db_xref="GI:5091627"
/translation="MSGNVNNOPLSGITIGSKSYVEQIWKIAGPLDIPKGE
KYTFPGHIEVLSLSLPLPSFSLHFSLSLSLSTETRELBELDIPICDPS
KLQGRVLAHLKVENNSDEYTABITLMPDTQVVLPTQENQFRLVNSFTKVLASD
TSAGGEFVPRKKAIECLPSLDSQPLAOELALIDLQGNQFRHNHNGTPORLLVLT
TGNNAFTISKLVAGDIVIVRGETELRVRIRARHQGNLPSISVSDCMRGVVA
SAKAFDNOCKFEYVYKPRSSKFIYSDELIDAVNNKFNVGSRFTMRLEDGSESRK
EGTIIYVDSFPHKCESEMSLEWODEFSFPGPKVSPMDIEHMLAVNPSFEL
KNKRLRYNETGSSSHLPLPITLQGOENQSVASDPNLSIRYDATEDAMNPSKIL
MSYVQPMKNTNNQVTEMEENITTKGTNRLRGYTLDPVYIKDPIEIGSEIS
KLTEGKRFGLSOTLRSPTEIQNKQFSSSRCTKVQOGVITGRADVLSNLTDLIL
ELEKLFIDKGLQTRNQWELAFITDSDKLVGDWPFCNMVKRIFIQKR"
complement(28888. .31377)
/gene="F23M19.5"
complement(28888. .31377)
/gene="F23M19.5"
/note="Contains similarity to gi|47935 protein kinase PKI
from Zea mays, is a member of the PF100954 S-100 locus
glycoprotein family and contains a PF10069 Eukaryotic
protein kinase domain."
/codon_start=1
/evidence=not_experimental
/protein_id="A039605.1"
/db_xref="GI:5091617"
/translation="MAVTPFKILPILLLHFPSPSTIGSVIYASGNOWPS
FNSTSVSFPSPNSFLAVSFAGSPYMSGYDSFNDPTDITVQSONVETGKILRS
TWDSKRIDRLVTSIEDTGEFTLLNNRSPVWSSFDNPTDITVQSONVETGKILRS
GLYSPOLERSGNLTLRNNTSAIYWNHNSFSFNSRLSLQNGVNSIYESWILG
GAEIYSGDYSNTFRELKILDDGNLRLTYSRSGSVNLAHWSVQCLVYGCN
FGICYNDFNPIICSPRNFEDVNDNRKGRKRYELDDCGNTMLDLVTRLEFY
HDDPSESEFAGSSPCRANCLSYLCLASVMSGDNOMKHGSGFEFGYQWPVS
TSYVYCGPVYANLRLRATRGDNNNSVHIVAVVAVIGILGAVAVETGLMCCCR
NPRGTESSHTLLEIYASCAPVOFTTELDORCTKSKREKLAGGREGTYRGVLYNTRY
VAVKOLEIEGEEKQFMEVATVISTHLLNLVLQFCQGRRLRVIEFMKNSLDN
PLEFTDSAKFLTWBYRNIALGAKGITYLHECRQCIYHCKIKPELILVDNFAKY
SLFGLTAKLLNDRNRYNMSVSRGRGLYAEMLANPIITKSIYSYGVALLVLESGK
RNFDVSEKTNKRKFSIMAYEEFEGNKALILDRLIEDDTVMEDQVMRAVKSFCIQ
EOPLOPRTMGVYVOMLEGITEIKNPDLKTIISVPSVSGMSMYSNASHMVAASGPTRSS
SSSARSPQMGITSSGSPASRTISEGMSGS"
complement(36027. .36827)
/gene="F23M19.6"
complement(36027. .36827)
/note="Contains two PF100560 leucine-rich repeat domains."
/codon_start=1
/evidence=not_experimental
/protein_id="A039606.1"
/db_xref="GI:5091618"
/translation="MINRRLVFCVAVVVSRLPYLAIEQVDPYKIVLPYVGR
FDPEEFTVSMOGNPCWENFGNCLGIIISISLNLIGTISPHRADLTSRLVDS
HNRKCTIPEETIKLKNLTIYDVSYNQLEHEVRVIRGAYLREPRADLEAVTLQVSP
TNKKRPVILVLLGIIVGVAGGAFEGYVLRIRKQPRLOEPRAEVATLQOQSSD
ESLVDESVYSIQIQRYLRRFWSKGLLLTLRQLKTNQNHLPYM"
40313. .42256
/gene="F23M19.7"
join(40313. .40937,41137. .41393,41705. .42256)

```

/gene="F23M19.7"
/codon_start=1
/evidence=not_experimental
/product="F23M19.7"
/protein_id="AAD39607.1"
/db_xref="GI:5091619"
/translation="MVPKIASMASRPLCSIPISPLSFLLFVYSLFEFSNLSNP
NPSISHTLONGINVVALEPRSLNYGLIDYSSSTPDSIPDPDPHPTKTSPPK
VPPENPLIKOYSAEVMMDLERSPEKRGSPAKRPFSSSDADVFPFRALSLSE
MELNGKSPFRKSGNEDYQORROYLDVKTITKARKSNGDHVLTDPVAMHVE
ETALSILVVDGSMFKODSSKNGSTLPERIQATQVSVINDVPTHLPLRDLISQ
NGRRSLTYFKAKHRHGGILREKLMDLVNEPQVMEDEFPNATGEOSIRMSRP
EFCILHAPGTPSCRLFDALISICIPVTSDTIELFEGDIIDYSEFVYSVASALP
KWLAMHGFSESEKELRSRIAKVOSFYVNDGNHADIIGIIEPNGANVHMKVQOK
VYMKAEVIRERERKPGASVPLRCCTI"
complement(43028..48860)
/gene="F23M19.8"
/complement(101n(43028..43288,43400..43672,43778..45067,
45156..45311,45523..45649,45915..46249,46361..47119,
47284..47637,47725..47869,48023..48613,48703..48736,
48815..48860))
/gene="F23M19.8"
/Note="Contains similarity to gi|36774 FABI protein from
Saccharomyces cerevisiae genome gb|D50617."
/codon_start=1
/evidence=not_experimental
/protein_id="AAD39608.1"

Query Match      5.7%; Score 383.2; DB 8; Length 88401;
Best Local Similarity 65.8%; Pred. No. 1.5e-56;
Matches 625; Conservative 0; Mismatches 258; Indels 67; Gaps 2;

QY 5732 ttattcaagttgtcgtcctcgaattttgacgtacatgtgatgcttcttca 5791
DB 72330 TTGTTTTGATCTTATCCATTTTAAACTTCCATACATTTGATGAGCTTTCTGA 72871
QY 5792 tcaatgtagaagcgtcgaatcagaacctccctgattgccaactgagagagatt 5851
DB 72870 ATGATTACAAAGCTCCACCTCACAGTTCCTCTAGCCTGGCATATAAGACGAATTC 72811
QY 5852 gcaactagatccttcaggagccatctaattgacgacatgtgacccaagattac 5911
DB 72810 GCGCTAGAGATCAGCAGAGGGTTGTCTTATCTTCAATGATCATTTGCCACCCAAATATAT 72751
QY 5912 catcgcgatgtaaaagctgcaaatatattatgtacgaaagtattgaggttctgagt 5971
DB 72750 CACCGTATGTGAAGCTGCTATATCTGTGTGACGAGGAATTTGAGCGGAGTAAAGCT 72691
QY 5972 gatttgggttagctggtcgaatgattacaagataccatggtacgactgtgaag 6031
DB 72690 GATTTCGGGTACTGACTATATGACTATMAAGATCACTCATGTCCAAAGGCTGTGCGT 72631
QY 6032 gtaaccatgtggacataagctccgagatccttcgacgtggaagtcatcagagaagacc 6091
DB 72630 GGGCATTTATGGACATATGCTCTGAGTACTCTCACTACAGGAAATCTTCAAGAAACT 72571
QY 6092 gatgtcttgggttagtggaataatgtctcctagagctcaattactgagcagggctttgat 6151
DB 72570 GATGTTTTTGGCTACGCGATCATGCTTTTGGAACTGATTAACAGAGTCAGAGACTTTTGAT 72511
QY 6152 ctgtcgtcgtcgaagcgtgtgattgtattgttggattgggtatggttccgggt 6211
DB 72510 CTGTGCAAGACTGGCGAATGCGATGACGATTAATCTCTCTAGATTGGGTA----- 72462
QY 6212 gtcccttggtaattattacataattagtgcttactacttggttggcccttggtt 6271
DB 72462 -----TAAACAAGATCTT 72450
QY 6272 ttatttcgtcctgattattccttagtcaatgttatgcatattgacctgtcttgcagt 6331
DB 72449 TTAGCAATATCTGGCTATCTCAAAAGCTGATTAATCTGTTCAATTGGGCT----- 72395
QY 6332 tcctttagtaaaagcctttagaagagaagaagcttgagatgctgctgcgtccctgacc 6391

```

```

DB 72395 --TCACAGTGAAGGGCTTTTGAAGACAGAACTGAGATGCTTGTGATCTCTGACC 72338
QY 6392 tgcagaacaattacattgacacagaagtgtgacgacttattcaagtagcatctctgta 6451
DB 72337 TGCAAACATATTACACAGAAAGCAAGATGAAACAGCTCATCAAGTGCCTCTTCTGCA 72278
QY 6452 cccaaggttgcgaatggagcgagccctaagaatgctcgaagtagtccgaatgcttgaagtg 6511
DB 72277 CACGAGACTCACTTAGAAGACGACCTTAAGATGCTTGAGTTTTCGAATGCTTGAAGTG 72218
QY 6512 atgaccttgcagaagaagtggagcagtggtgcaaaagttaagattcatccatcaagacgtag 6571
DB 72217 ACGGTTAGCGGAGAAATGGAGCAAGTGGCAGAAAGTGAATCTTCAGCGCAAGAGTGG 72158
QY 6572 aattagctccaatcgaactctgaaatgatactccttagactcgacagataacttgatctt 6631
DB 72157 AGCTCTCTCTCACCCACCTCTGACTGATCTTGATTCGACTATATCTTCATGCTTA 72098
QY 6632 ttgaattatcgtgtccagaataacagcatataaagttaatgaatataa 6681
DB 72097 TGAAGTTGTGTGCTCCAGATTAACGACATTTGATTTGCTTAACAGANA 72048

```

RESULT 6

```

AC012654 98471 bp DNA PIN 10-DEC-1999
LOCUS AC012654
DEFINITION Arabidopsis thaliana chromosome 1 BAC F14023 sequence, complete
sequence.
ACCESSION AC012654
VERSION AC012654.2 GI:6554462
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

```

REFERENCE

1 (bases 1 to 98471)
Liu, S.X., Lee, J.M., Sakano, H., Yu, G., Jhanveri, A., Lenz, C.,

Tortumi, M., Chin, C., Chiu, J., Choi, E., Gonzalez, A., Howing, B.,
Koo, T., Li, J., Liu, A., Pham, P., Vaysberg, M., Altatli, H., Brooks, S.,
Buehler, E., Chao, Q., Conn, L., Conway, A. B., Hansen, N.,
Johnson-Hopson, C., Khan, S., Kim, C., Lam, B., Nguyen, M., Palm, C.,
Shin, P., Tambunga, G., Davis, R.W., Ecker, J.R., Federspiel, N.A. and
Theologis, A.

Arabidopsis thaliana chromosome 1 BAC F14023 sequence

JOURNAL Unpublished
2 (bases 1 to 98471)

REFERENCE Theologis, A.

AUTHORS Direct Submission
Submitted (02-NOV-1999) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA

REFERENCE 3 (bases 1 to 98471)

AUTHORS Theologis, A.
Direct Submission
Submitted (10-DEC-1999) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA

JOURNAL On Dec 10, 1999 this sequence version replaced gi:6151511.

COMMENT location/Qualifiers

FEATURES

1..98471
/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="1"

/clone="F14023"

BASE COUNT 31501 a 17757 c 17475 g 31738 t

ORIGIN

```

Query Match      5.7%; Score 378.6; DB 8; Length 98471;
Best Local Similarity 58.7%; Pred. No. 8.7e-56;
Matches 863; Conservative 0; Mismatches 454; Indels 152; Gaps 6;

```


QY	4924	tacgtcttggaactctgtttcttcacgtcttagtctaactcaagctgaagagaccaga	4983
Db	79817	CAAACTCTTTGGAGTAATAATTCGATTGCAGAAATCTTTTGCAACCCAGAGAGATCCAGA	79876
QY	4984	agtcacaccttgctcaactgaagaagtttctctctgcagaaattgcgaagtcgcaacgatalc	5043
Db	79877	AGTTCATCTGGGAGACGTCAAGAGGTTTCTTTCTGGGAGCTACCAAGTGGGACTGATGG	79936
QY	5044	ttttagt-----acatccttggaaagagtgatgttggtaaggtgtatataaggaagcct	5097
Db	79937	GTTTAGTAACAAGAACATTTTGGGCAAGGCGGTTTGGGAAAGCTTCAAGGGAGCCTT	79996
QY	5098	tgctgatgctcaccattcttagcagcttaaaagccttaaaagaacgaacacccaggtggcga	5157
Db	79997	GGCAGACGGAGACTTTGTGTCTGTCCAAAGACCTAAGAGAAAGCGAACTCCAGGTGGAGA	80056
QY	5158	gctcgacgttccaacagaagaatggaatgatagcatgctgtgcatcgaaatctcttcgcg	5217
Db	80057	GCTCCAGTTTCAAAACAGAAAGTAGATGATGATGATGCAATTCATCGAAACCTGTTGAG	80116
QY	5218	tctacgtggttcttcgcatagcactaccgagcgcctctcttatatcatcatcagctcaa	5277
Db	80117	ATTACGAGAGTTTCTGTATAGACACCGACGAAAGATGCTGTGTATCTTACATGACGCCAA	80176
QY	5278	tggaagtgtgcgtcatgtttaagaagta	5306
Db	80177	TGGAAAGTTCTCTCGTCTCAGACAGTA	80205
RESULT	7		
ATF17M5/c			
LOCUS	ATF17M5	96475 bp	DNA
DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone F17M5 (ESSA		PLN 11-MAR-1999
ACCESSION	AL035678.1	GI:4490291	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
	thale cress.		
	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons;		
	Rosidae; Caprariales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 96475)		
	Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Hohensei, J.,		
	Mewes, H.W., Mannhaupt, G., Mayer, K.F.X. and Schueller, C.		
	Unpublished		
	2 (bases 1 to 96475)		
	EU Arabidopsis sequencing, project.		
	Direct Submission		
	Submitted (11-MAR-1999) MIPS, at the Max-Planck-Institut fuer		
	Biochemie, Am Klopferspitze 18a, D-82152 Martinsried, FRG, E-mail:		
	schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project		
	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge		
	Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,		
	E-mail: michael.bevan@bsrc.ac.uk		
	Information on performance of analysis and a more detailed		
	annotation of this entry and other sequences of chromosome 4 can be		
	viewed at: http://websites.mips.biochem.mpg.de/proj/thal/ .		
FEATURES			
Source			
	1..96475		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
	2377..3928		
	/gene="F17M5.10"		
	complement(2377..2571)		
	/number=1		
	complement(join(2377..2571,2755..2837,3032..3098,		
	3180..3230,3461..3574,3668..3724,3815..3928))		
	/note="Similarity to Caenorhabditis elegans com1d, gene		
	T1661.11, FID:e1369366; contains EST gb:120620, N65143,		

/codon_start=1
/product="putative protein"
/protein_id="CAB38783.1"
/db_xref="GI:4490292"
/translation="MGVEIOSPOSSVYVEQVLAALNPENELIPDENYVNTYSOTY
SLEVNLCILRIYKOPFERMNTIYARILYALMAMPDPDSICFLIPEVOMERPK
SLIVSHLYETGEREQQWDEAKNRHILEANPGEQALQAYASHLSLSLOKIPRSVL
AEAVNMDSASIDKFLTEQVINSGMIVKEGGSIVLPONEFNHPELKNKTGENVPLEHI
ARIPILG"
/number=1
complement(2572. .2754)
/number=1
complement(2755. .2837)
/number=2
complement(2838. .3031)
/number=2
complement(3032. .3098)
/number=3
complement(3099. .3179)
/number=3
complement(3180. .3230)
/number=4
complement(3231. .3460)
/number=4
complement(3461. .3574)
/number=5
complement(3575. .3667)
/number=5
complement(3668. .3724)
/number=6
complement(3725. .3814)
/number=6
complement(3815. .3928)
/number=7
4544. .6491
/gene="F17M5.20"
complement(4544. .4687)
/number=1
complement(join(4544. .4687,4781. .5362,5443. .5712,
5815. .6093,6423. .6491))
/note="strong similarity to WD-repeat protein -Daucus
carota, PID:g2253631: Contains trp-asp (WD-40) repeats
signature [VLAIALDHTVYLMDA] [TGGGGGDRITKFWNT]
[VASAGDETLRFMNV]"
/codon_start=1
/product="WD-repeat protein-like protein"
/protein_id="CAB38784.1"
/db_xref="GI:4490293"
/translation="MDAGLNRCPLOEHLPRKNSKENIDRFIPNRSAMNPDYAFALT
EERKGDQSAIVSSPSKEATRKQLAETMLNHTRIATRNKPQAPVELLPNSHSLH
QQPKSVKPRRTIPQTSERTLDAPDIVDFYLNLDWGSANVLAIALDHTVYLMDASG
SISELVTIDEKGPVYSINMAPDGRHVAVGLNNSVQLMDSASNRQLATLKGQSHV
GSLAMNHILITGGMDGLIINDVRIRSPIVETVYRGHTQVCGKMSGCGGLASGCGN
DNVYHITWRSVASSNSTOTQWLHRLHEHTSAVKALMCPQOANLATGGGGDRITKFM
NHTTGACLSNPDIGSOVCSILSKNERELLSHGFTONQLTLMVYPSVYKAEITLGT
SRVLYMAQSPDGCYVASAGDETLRFMNVGVPETAKKAAKAVAEPSVNRIR"
complement(4688. .4780)
/number=1
complement(4781. .5362)
/number=2
complement(5363. .5442)
/number=2
complement(5443. .5712)
/number=3
complement(5713. .5814)
/number=3
complement(5815. .6093)
/number=4
complement(6094. .6422)
/number=4
complement(6423. .6491)
/number=5
7856. .9901
/gene="F17M5.30"

exon
complement(7856. .7999)
/number=1
complement(join(7856. .7999,8104. .8685,8767. .9036,
9139. .9417,9803. .9901))
/note="strong similarity to WD-repeat protein -Daucus
carota, PID:g2253631: Contains trp-asp (WP-40) repeats
signature [VLAIALDHTVYLMDA] [TGGGGGDRITKFWNT]
[VASAGDETLRFMNV]"
/codon_start=1
/product="WD-repeat protein-like protein"
/protein_id="CAB38785.1"
/db_xref="GI:4490294"
/translation="MDAGNNTSSHYKTOARCPLOEHLPRKNSKENIDRFIPNRSAM
NPDYAFALTGKRGKQDOTAAVSSPSKEATRKQLATRNALNHTRIATRNKPQAVEL
LPNSHSASLHQPKSVKPRRTIPQTSERTLDAPDIVDFYLNLDWGSANVLAIALDH
TVYLMDSAGTSELTIDEKGPVYSINMAPDGRHVAVGLNNSVQLMDSASNRQLR
TLKGHQS RVGSLAMNHILITGGMDGLIINDVRIRSPIVETVYRGHTQVCGKMSG
SGOOLASGDNVYHITWRSVASSNSTOTQWLHRLHEHTSAVKALMCPQOANLATGG
GGDRITKFWNTHTGACLSNPDIGSOVCSILSKNERELLSHGFTONQLTLMVYPSM
VKMAELTGHTRSVLYMAQSPDGCYVASAGDETLRFMNVGVPETAKKAAKAVSEPF
SHVNRIR"
/number=1
complement(8000. .8103)
/number=1
complement(8104. .8685)
/number=2
complement(8686. .8766)
/number=2
complement(8767. .9036)
/number=3
complement(9037. .9138)
/number=3
complement(9139. .9417)
/number=4
complement(9418. .9802)
/number=4
complement(9803. .9901)
/number=5
complement(10668. .10967)
/number=1
10668. .12621
/gene="F17M5.40"
complement(join(10668. .10967,11056. .11159,11587. .12457,
12511. .12621))
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB38786.1"
/db_xref="GI:4490295"
/translation="MEENCDECMKREBELYWTHTQTLHPTOLLPGFHNRLQPMLLFO
FLVYTFICFPAPRELEHTYTQTORGLSNFCFFLSKSVLCGKSPKIANSSITITFC
NLLOVILPKRFSTHCKRKLPOVTLKSPSGVYVNGVSEDEKTMARFEGDKRVKSHS
LEENDLVFKHGVSEEVLEVDQTLCEKPTSYFVBCGHAERTKGIIDFNATSSRS
PEHNNDPADVEPTPNOQLVISPVDNLEEDLIDIDFDIDKILNPLVASHGYROE
EIRINDITDQLOPVLISPTSTVAVSGKYPLDSGFKMRRELSDNINDOAKASNRKA
LSIARASISPGFLVEMKRSVYKCFLLTIPKGCYKVMNLTITQEVYMOVDQKMKR
NIRGARSGGISTGMKFTVQDNNLREGDVCVPEPANSERKPLHINVIIRGSETIRT
NNDVYITISE"
complement(10968. .11055)
/number=1
complement(11056. .11159)
/number=2
complement(11160. .11586)
/number=2
complement(11587. .12457)
/number=3
complement(12458. .12510)
/number=3
complement(12511. .12621)
/number=4
12884. .14176
/gene="F17M5.50"
12884. .14176
/gene="F17M5.50"
/note="strong similarity to Arabidopsis thaliana

Query Match	4.7%; Score 317.6; DB 7; Length 96475;	chromosome II BAC (gene T6A23.21), PID:g3786013"
Best Local Similarity	66.2%; Pred. No. 2.4e-45;	/codon_start=1
Matches 522; Conservative 0; Mismatches 204; Indels 62; Gaps 2;		/product="hypothetical protein"
		/protein_id="CAB38787.1"
		/db_xref="GI:4490296"
		/translation="MTDPLKDLIEIILSRVSMSTRVRLTRCKSNLTSSEPRKM
		HIGKATSTREGRVIMLDYFLFMSAVINDNDPSTLEFGKISLCEQVKISQVH
		CSGLICITKEDTRIVMNPRIQOETRWIIPRSHRPYVIMNLSVQSLGNKSWLCA
		LRFLDYCTEKHICMHEIYDSDLTWTLDTAPHWYLLSNMCSVQSLGNKSWLCA
		REBNDGYNHIIICFDFTRERFGLPLPLPNVINDNEEYVTSVCEGKIALFQHNDS
		YVELEIWTTRIEAEVMSNMFRLIDIEPNINIAVPEFYGGFFIDEKKKVALGTFDE
		EFGRKTEINIGEDGYFEREDRITFINIEBAGRAGVNCSTVCSVPELVIRIKRPAGC
QY 5799	tagagcgtagcagccatgaacaccccccattgagccaaatagggagagattgacag	5858
DB 50766	TAGAGCTCCCGAGTCCACACCACTTGATGGCCAAAGAGACGCTATTCCTTG	50707
QY 5859	gattcttaggggacctaaattgaatgacatgacattgacccaaatattccatgcg	5918
DB 50706	GATCTGCAGAGGGCTGCGTATTTACATGATGCGACCCAAAGATTATTCATCGAG	50647
QY 5919	atgtaaaagctgcaaatattatttggacgaagaatttggagctgtttagtgatttg	5978
DB 50646	ATGTGAAGCTGCAAAATTTTGTGGATGAAGATTGTAAGCCGGTGTGGGATTTG	50587
QY 5979	gattagtagcctatgataacaggaataccattgtaacagctgtgtaagggtatca	6038
DB 50586	GACTTGCAAAACATGATGATACAAAGACACATGACACCGAGTGGCGGACAA	50527
QY 6039	ttggagccatagctcccgatgacctctgcagctggaaaagtcatcagaagaaccg	6098
DB 50526	TTGGTCAATAGCCCTGAGTACCTTTCCACTGGAATATCATGAGAAACGAGTGT	50467
QY 6099	ttgtttatggataatgctcctcctagagctcattcctgacagagagctttgactgc	6158
DB 50466	TTGGGTATGGAGTCATCTCTTGAGCTTATCACTGACAAAGGCTTTGATCTTG	50407
QY 6159	gccttcgcaacagatgatatgttgttgattggatgtagtgcgggtgttcctt	6218
DB 50406	GCCCTCGGAGATGATGATGATGTCTTACTAGACTGGTATAT-----	50362
QY 6219	tgttaattattcacaattagtgcttactactgtgttgcccttgtttatttc	6278
DB 50362	-----ACATATATAGTTCCTCCACTT-----	50341
QY 6279	ctgcctgatttgatctctatgcatgcatatgacatgcttgcgaatgcttta	6338
DB 50341	-----GTATAGTTCATATCTGTGTACCTCATTTGTGTAATATTCATTTCGACA	50289
QY 6339	ggttaaaagcttggaaaagaaaagttggagatgctgtgcgactccatgacagaa	6398
DB 50288	GGTAAAGGGTTGTTAAAGAGAAATTTGGAAGCACTAGTAGATGTTGATTTTCA	50229
QY 6399	caattacatgacacagaagtgtgagcagcttatcgaagtacgacttactcgttacc	6458
DB 50228	TAAATTACAAAGACAGAGAGTGGAGCACTAATCCAGTGGCTTTACTCTGCACTGAG	50169
QY 6459	ttgcgcaatggagcgccatgaatgtagaggttagtccgaatccttgaagggtatgc	6518
DB 50168	TTTACACATATGGAAGACCCCAAAATGCTGTAAGTGTAGAGATGCTTGGAAGA	50109
QY 6519	ttgcgaaagtgagcagtggtgcaaaaagtgaatcatcatcaagaagtagaatagc	6578
DB 50108	AGCTGAGAGATGGAGAGATGGCAAAAGAGAGAAATGTTACAGACAGATTTCAC	50049
QY 6579	ttccacatc 6586	
DB 50048	AACCCACC 50041	

RESULT	8
AC006436	
LOCUS	
DEFINITION	Arabidopsis thaliana chromosome II section 77 of 255 of the complete sequence.
ACCESSION	AC006436 AE002093
VERSION	AC006436.4 GI:6598565
KEYWORDS	HTG.
SOURCE	
ORGANISM	thale cress.
REFERENCE	Arabidopsis thaliana
ADTHORS	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta; eudicotyledons: Magnoliopsida: Euphorbiales: Euphorbiaceae: core eudicots: Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsi.
ADTHORS	1 (bases 1 to 65899)
ADTHORS	Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldhahn, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Renning, C.M., Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., Vanden, S.E., Umayam, L., Tallon, L.J., Gail, J.E., Adams, M.D., Carrera, A.V., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D., Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M., and Venter, J.C.
TITLE	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
JOURNAL	Nature 402 (6763), 761-768 (1999)
MEDLINE	20083487
REFERENCE	2 (bases 1 to 65899)
ADTHORS	Lin, X.
TITLE	Direct Submission
JOURNAL	Submitted (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT	On Dec 17, 1999 this sequence version replaced gi:4726109. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (http://www.tigr.org/tldb/at/at.html).

Genes were identified by a combination of three methods: Gene prediction programs including GRLI (<http://artur.eprn.ornl.gov/pub/xyrail>), GeneFinder (Phil Green, University of Washington), GENSCAN (Chris Burge, <http://genome.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curation based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to the bottom of the chromosome.

We thank the GSHL/Mashu/ABI consortium for sequencing BAC clones F6P23, F5U6, T17A5, and T13J16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Juming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khataf, Michael E. Heaney, Lily Fu, Feng Jiang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: atltigr.org.

FEATURES		Location/Qualifiers
source		1..65899 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="2" 1..65698
misc_feature		/note="Sequence from clone F13J11" complement(join(451..842,907..1234)) /gene="At2g13660"
mRNA		complement(451..1234) /gene="At2g13660"
gene		/note="F13J11.1" complement(join(517..842,907..1114)) /gene="At2g13660"
CDS		/note="unknown protein" /codon_start=1 /protein_id="AAD28310.1" /db_xref="GI:4726110"
repeat_region		/translation="MCCGMLREKGFANTDSSPEPTFISCHRLITNGNFIHQPADLSPRNLYLCLRLYLAFCHRSPEETEDSDNSSTVPIKINILINPRTDALNKFWSKELPSLGRRIKHSNMKDPDFWNHFLPEIGRVNLSLOKATRKRRKRVNPKSERETGKGNLKHKKSSD"
mRNA		complement(2606..2687) /rpt_family="GAA)n"
gene		join(<3349..3541,3644..3744,4109..4230,4375..4471,4583..4693,4792..4928,5629..5755,5848..5997) /gene="At2g13670"
CDS		/gene="At2g13670" /note="F13J11.2; predicted by genscan and genefinder" join(3349..3541,3644..3744,4109..4230,4375..4471,4583..4693,4792..4928,5629..5755,5848..5997) /gene="At2g13670"
repeat_region		complement(2606..2687) /rpt_family="GAA)n"
mRNA		join(<3349..3541,3644..3744,4109..4230,4375..4471,4583..4693,4792..4928,5629..5755,5848..5997) /gene="At2g13670"
gene		/note="hypothetical protein" /codon_start=1 /protein_id="AAD28311.1" /db_xref="GI:4726111"
CDS		/translation="MAOSTSHDSGPOGLMRRPSSAATVTSIEVFHEVAVSLGTLAPLRYAAIEHEHRRPVATLCRFVAFKRAHRLDPSGGVROFKTLIFQELMANS SLASRYKTDGREVEFYQOYEHYVRLDQDADRAOLGKAYCTAGVLEPYTMAN KSEKVAVAPETIIAARDYQENETIYAPNITIPDSASQSVQLEEVYKAAVALN TRGLNMPSGEBOHRTGMJLDLWLRAFEGQASWLPQAAODIQKRLTNGLYLLI WGEAANIIRMPSCLCYIFHNMAVELHGLAGNVSIVGENIKPSYGGDEAFLRKVT PIYRVQTVQV"
gene		10367..13711 /gene="At2g13680"
mRNA		/note="F13J11.3"
gene		/pseudo complement(join(<14929..15843,16401..>17114)) /gene="At2g13690"
CDS		complement(<14929..>17114) /gene="At2g13690"
gene		/note="F13J11.4; predicted by genscan and genefinder" complement(join(14929..15843,16401..17114)) /gene="At2g13690"
mRNA		/note="hypothetical protein" /codon_start=1 /protein_id="AAD28312.1" /db_xref="GI:4726112"
gene		/translation="MGSDSLRKPTNGTGLSRRRSWCCSFAPVPPSPDTRISRNH IPAKSOQREKLVPSCPOSSKALNVNIDRRILISGVRSPIDSDPTVMDETER TOBEDDAVYVDTPNLRSRSPAKITVYTGSLSGVDARLSLGRNGGVILVLETS LBYLANAVYVSGSLINEKRCSSSSSLKATCTIETCDVENCLOVPEYELMIEES NVILKRMIVGVRALVDLEAAGIFSRVAVJSCVLYEAVPWTEDEERIKRLIGIT SFDDAVSEILKEDIVRAGKCVDSLAKLEBSSSSSSSKKEPLLESLSREVINML LEIMIDREIAEEFVEMIGKORRLVHEHNSVMEVYEVSRVGAFFIANGRRVQCG EARAGVLAEMFELVDFGLORCKGLDMREVEGMSGQTLTLTVKQYQVEMEMER WFSKHGTECPNLKAFQIWMRSRFLRGVSSSCR"
gene		complement(18082..20371) /gene="At2g13700"
repeat_region		complement(18082..20371) /note="F13J11.5"
repeat_region		/pseudo 29450..29498 /rpt_family="POLY_A"
gene		complement(29932..30854) /gene="At2g13710"
repeat_region		/note="F13J11.6"
repeat_region		/pseudo 31008..31094 /rpt_family="GAA)n"
repeat_region		complement(31832..31913) /rpt_family="GAA)n"
mRNA		join(32232..32373,32688..32771,32839..32939,33013..33167,33252..>33473) /gene="At2g13720"
gene		32232..>33473 /gene="At2g13720"
CDS		/note="F13J11.7" join(32264..32373,32688..32771,32839..32939,33013..33167,33252..33473) /gene="At2g13720"
repeat_region		/note="unknown protein" /codon_start=1 /protein_id="AAD28313.1"
repeat_region		/db_xref="GI:4726113"
mRNA		/translation="MVALVAQERLGSNRCASRMDSRMPGRMPPRVPRDGLKSPITV LVTPYKRSVTFSSSHKFFYTLNLSVYTSVTLNTRGKKDKDRNLTIKRYKGVW QVREKRNLSVNDQNKRFITNDSSEALNNAFGTSDKOKNTEVTEGTLDSSKQ QIMVSDVFTKRLKLYFTTSNQRHININKNNTILKNDKSKNIIILKKYKVVLYII IVLI"
repeat_region		33343..33441 /rpt_family="POLY_A"
gene		complement(<33782..>34228) /gene="At2g13730"
CDS		complement(<33782..>34228) /note="F13J11.8"
repeat_region		complement(33782..34228) /gene="At2g13730"
repeat_region		/note="unknown protein" /codon_start=1 /protein_id="AAD28314.1"
repeat_region		/db_xref="GI:4726114"
repeat_region		/translation="WADVIRADAKGYMREVEDQAVNEAGQRLDDHALIPEDANDQ ARLNQLAARDAAGLIVLDLHNRGIDRLHWNQNDALGVENHNNGVRHOPQIPAAITP LAEPRRTIGDFNRPQGMFYANRSALVPPPYORNDYESHLYITLHI"
repeat_region		complement(35689..35776) /rpt_family="TAAAA)n"
gene		36924..37497 /gene="At2g13740"
repeat_region		/note="F13J11.9"
repeat_region		/pseudo complement(37457..37522) /rpt_family="GAAA)n"
mRNA		join(<38993..39051,39561..39686,39797..39989,40095..>40208) /gene="At2g13750"
gene		<38993..>40208
Query Match		4.7%; Score 313.6; DB 50; Length 65899;
Best Local Similarity		64.8%; Pred. No. 1.2e-44;
Matches 542; Conservative		0; Mismatches 239; Indels 55; Gaps 3;
QY	5716	taatgctccacttaattcaatgctgctcctcctgaattttgatctgacatg 5775
DB	64398	TATTGTCATATATTTTGTGTTTGTACGTTGTGATTTGCTAGTTGTAAGTAT 64457
QY	5776	tgatctctgtttccatcaatgtagacgtcaagcatcaaacctccctgattg 5835
DB	64458	CTATATCTTTGATTT--TCATATATAGAGGCTCTGAAGCAATCAGCCTTGACTGCC 64515
QY	5836	aactaaggagaagatgcactagatcctctagggccatctaaattgcatac 5885
DB	64516	AAAAAGAAACATATTGCTCTGGGATCAACAAGGGGCTCGCATATTTTACAGATCATTG 64575

QY	5896	tggcccaagattatccatccgcgagatgtaaaagctcacaatattatattggacgaagatt	5955
Db	64576	CGACCAAAAGATATTACCTGGATGTGAAGGCTCCAAATTAATGTTAGATGAAGGTT	64635
QY	5956	tgaagctgtgtatgagtgatcttgggtttagctagctcattgatacaagaataccatgt	6015
Db	64636	TGAAGCTGTGTGGAGATTTTGGCGCTAGCAAAATTAATGATTAATACGACTGCCATGT	64695
QY	6016	taagactgcgttagagggtacacatttggagacataagctccgaatgactcttcagatgaaa	6075
Db	64696	GACACTGCTGTACGGGGAGACGATTTGGCCATATACGCCCGAGTACTCTTCGACAGGAA	64755
QY	6076	gtcaatcagaagaacacgaatgctcttggatattggaataatgctcctagaagctcattg	6135
Db	64756	ATCTTGAGAAAGACATGATGTTTTTGGTACGGGGCTCATGCTTCGAGCTCATCACTGG	64815
QY	6136	acgaagagagcttttgatccttgccttcgcttcgagacgatataatgtaattgttggatg	6195
Db	64816	ACAAAAGGCTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	64875
QY	6196	gtaatgctcccgagtgcttccttggatattatcattcacaataatgagtgcttactactt	6255
Db	64876	GGTATGTTTTTAATAGTCCCGATCATATTTTGAATCAAAACTGT-----TATTTTG	64929
QY	6256	tttgagcccttgtttttatttcttcctgctgtatttgattccttagcatgtaicatatc	6315
Db	64930	CAATAAAGAGAGAGTCTTTCTTCTGAGTGA-----	64963
QY	6316	gacctgtcttgcgaatgcttctttagtgtaaaagccttttgaagaagaaagttagagatg	6375
Db	64963	-----TAGTGAAAGAGGTTTGAAAGAGAAATTGGAAGC	65002
QY	6376	ctgtgcatcctgaactgcgcgaacaattacattgacacagaagttgagcaactattca	6435
Db	65003	CTTGTGATGCAGAACTCGAAGAAAGTCACTGGAACAGAAAGTGGACGACTGATACAA	65062
QY	6436	gtacgacttctcttaccagaggtctgcacatgagagcgagcctaagatgtcagagtgatgc	6495
Db	65063	ATGGCTTGCTGCTGCACTCAAAAGTTTGCATATGGAAGACGCCAAAGATGTCAAGAGT	65122
QY	6496	cgaaatgcttgaagtgatgagccttgcagaagaatgagagagtgagtcgcaaaagtcca	6551
Db	65123	AGAACTGCGAAGAGATGTTAGCTGAGAGATGGGAAGATGGCAAAAGAGAGA	65178
RESULT	9		
LOCUS	A67827	2089 bp	DNA
DEFINITION	Sequence 32 from Patent WO9743427.	PAT	05-MAY-1999
ACCESSION	A67827		
VERSION	A67827.1	GI:4756649	
KEYWORDS			
SOURCE			
ORGANISM			
			thale cress.
			Arabidopsis thaliana
			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
			euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
			eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
			Arabidops.
REFERENCE			1 (bases 1 to 2089)
AUTHORS			De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE			PRODUCTION OF APOMICRITIC SEED
JOURNAL			Patent: WO 9743427-A 20-NOV-1997;
			CIBA GEIGY AG (CH)
FEATURES			
source			Location/Qualifiers
			1..2089
			/organism="Arabidopsis thaliana"
			/db_xref="taxon:3702"
			/clone="SERK GENE CDNA"
			195..2072
			/note="unnamed protein product"
			/codon_start=1
			/protein_id="CAB42254.1"
CDS			

BASE COUNT 569 a 419 c 502 g 599 t
ORIGIN

Query Match	Similarity	Score	250.4	DB	5	Length	2089
Matches	524	Conservative	0	Mismatches	201	Indels	143
							Gaps
QY	5800	agagcgtcagccatcagaaacctccccctgattgagccaaataggagagagatgacactagg	5859				
DB	1361	AGAGAGGCCACCGTCACCAACTCCGCTTGATGGCCAAACGGGAAGAAATCCCGCTAAG	1420				
QY	5860	atctctcaggagcctatcctaattgacagacattgacatggatcccaagattacatcgga	5919				
DB	1421	CTCAGCTCGAGGTTTGCTTACCTCATATGATATCAGCGCATCCGAAGATCAATTACCGTGA	1480				
QY	5920	tgtaaaagctgcaaatatattatgacagaaagattgagctgtgtgtaagtatttgg	5979				
DB	1481	CGTMAAAGCAGCAAACTCCTCTTAGAGAGAAATGCAACGGGTGTGGAGATTTCGG	1540				
QY	5980	gttagctagagctcattgattacaagagatcccatgttgaagctcgttgaagggatccat	6039				
DB	1541	GTTGGCAAACTTATGGACTATTAAGACACTCCTGTGCAACACCACTCGTGGCACAT	1600				
QY	6040	tgggacatatgctcccgagtaacctctgcagctgagaaagtcacagagagacagatctct	6099				
DB	1601	CGGTACATGCTCTCCAGATATATCTCAACCGGAAATCTTCAGAGAAACCGACGTTT	1660				
QY	6100	tggttatggagataatgctccctagagctcatattgacaagagagcctttgactctggc	6159				
DB	1661	CGGATAGGGAATATGCTTCTAGAACTATATACAGAGCAAGAGACTTTCGATCTCCGCTG	1720				
QY	6160	ccttgcgaaagatgagatgattgtgtgtgtgattgggtgatgtgccgggtgttcctt	6219				
DB	1721	GCTAGCTTAACGACACACACGCTCACTGTTACTTACTGACG	1757				
QY	6220	ggttaattatcacatatagtgcttactactctgttctgtgacctgttttatttcc	6279				
DB	1757	-----	1757				
QY	6280	tgccgttatgtattccttagtcagtattgacatatgacactgcttgcaatgtcctttag	6339				
DB	1757	-----	1757				
QY	6340	gttaaaagccctttgaaagaaagaaagtgtggatagtcgctgcagtcctgcagcagaa	6399				
DB	1758	GTTAAAGGATGTTGAAGGAGAAAGACTAGAGATGTTAAGTATGCCAGATCTTCAACA	1817				
QY	6400	aattacattgacacagaagtgtgacagcttattcaagtaagatattactgtaccaggt	6459				
DB	1818	AACCTCAGAGAGAGAACTGGAACAAGATACAAAGTGGGCTTGATGACGCCAAGA	1877				
QY	6460	tcggcaatggagcggcctaagatgtcaagsggtagtcocgaatgtcttgaagtgatgacct	6519				
DB	1878	TCACCAATGGAAACCAACCAAGATGCTGAAAGTGTATAGCATGCTGGAAGGAGATGGCTT	1937				
QY	6520	gtagaagaagtggcgctgggtgcaaaaatttgaagatcatcatcaagaagatgaattagct	6579				
DB	1938	GGGGAAGATGGGCGAATGGCAAAAAGTTGAGTTTGAAGGAAGAAATGATTTAGT	1997				

OY	6580	ccatcgaactctcgatgcatcctagacgcagaagtaacttgcagctttgaata	6639
Db	1998	cctaactcctaacttcgatgttgattcttcttgaatttcaatttgcagcccttagatta	2057
OY	6640	tctgtccaagataaacagcatataaa	6667
Db	2058	tctgtccaagataaaaaaaaaaaaaa	2085
RESULT	10		
ABO13395/c			
LOCUS	ABO13395	PLN	20-NOV-1999
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone: MON23,		
ACCESSION	ABO13395		
VERSION	ABO13395.1		
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (strain:columbia) DNA, clone_l1db:Mitsui P1		
ORGANISM	clone:MON23.		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Euhayriota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
JOURNAL	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
REFERENCE	Arabidopsis.		
AUTHORS	1 (sites).		
TITLE	Nakamura.Y.		
JOURNAL	Structural Analysis of Arabidopsis thaliana Chromosome 5. VI		
REFERENCE	Unpublished (1998)		
AUTHORS	2 (bases 1 to 86064)		
TITLE	Nakamura.Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (06-MAY-1998) to the DDBJ/EMBL/GenBank databases.		
AUTHORS	Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of		
TITLE	Gene Structure 2; I532-3, Yana, Kisarazu, Chiba 292, Japan		
REFERENCE	(E-mail:yakamukazusa.or.jp, Tel:+81-438-52-3935,		
AUTHORS	Fax:+81-438-52-3934)		
TITLE	Location/Qualifiers		
JOURNAL	1..86064		
REFERENCE	/organism="Arabidopsis thaliana"		
AUTHORS	/strain="Columbia"		
TITLE	/db_xref="taxon:3702"		
JOURNAL	/chromosome="5"		
REFERENCE	/clone="MON23"		
AUTHORS	/clone_l1db="Mitsui P1"		
TITLE	/BASE COUNT 27710 a 15313 c 14986 g 28055 t		
JOURNAL	ORIGIN		
Query Match	3.2%; Score 217; DB 7; Length 86064;		
Best Local Similarity	55.6%; Pred. No. 3.9e-28;		
Matches 513; Conservative	0; Mismatches 335; Indels 75; Gaps 2;		
OY	5751	cctcgaatttttgatcgtgaacattgtgagtcttctgtttc---atcaaatgtagaagcgtc	5807
Db	73777	CTATAAATCTTGGATTTCAGGAGTTTAAGAATTTTGATCTATATACATMGAGAGATT	73718
OY	5808	aggcatacaacctcccctgattggccaactcaggagagagattgcaactctcta	5867
Db	73717	AACC GGCGGATCCAGTTCGGAATGGTTTGAAGAGAAAGATGGGTTAGGTGACGCAC	73658
OY	5868	ggggcctactcaaatatgcatgaccattgtgatcccaagattcacatcgcgatgtaaag	5927
Db	73657	GAGGACTCAAAATATCTTCATGAACAATTGCCAACCCGAAGATCATCACAGAGATGAAAG	73598
OY	5928	ctgcaatatattatggcgcgaagaatttggagctgttgttagtgatttttggttgcca	5987
Db	73597	CTGCAAAATGTGTTACTACAGATGAAGACTTTGAACCACTGGTGGTGAATTTGGTTAGCCA	73538
OY	5988	ggctcattgataccaagagatcagctgttcagcgtctgtaaagggtacattgggcaca	6047
Db	73537	AGTGGTAAATGTTAAGAAGCACTAATGTATCAACCACTAGGTCCGAGAACAAATGGGTCA	73478
OY	6048	tagctccgaagtaacctcgtgactgtgaaaagtcatacagaagaagaccgattgtcttggtatg	6107

[illegible]

[illegible][illegible]

Db 108419 TGTGGCCACATTGGCAGCTGAGTACTTATCCAGCGGTCACTGCTCAGAGAGAGTATGT 108360
QY 6097 ctttggtatgagataatgctccttagatcattactgacagaggcttttattctgc 6156
Db 108359 CTTGGCTTTGGCATCTCTCTTGTAGCTCATGTGTCAGAAAGCTCTGATTTGG 108300
QY 6157 tcgccttgacagatgatgtatgtt 6186
Db 108299 CAGATCCGCACACCGAAGGTATGCT 108270

RESULT 13
LOCUS F17123 134784 bp DNA PLN 20-JUN-1999
DEFINITION Arabidopsis thaliana BAC F17123.
ACCESSION AF160182
VERSION AF160182.1 GI:5106766
KEYWORDS
SOURCE .
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsids.
REFERENCE 1 (bases 1 to 134784)
AUTHORS Washington University Genome Sequencing Center.
TITLE The A. thaliana Genome Sequencing Project
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 134784)
AUTHORS Sun, H., Wohldmann, P., Johnson, D. and Gibson, A.
TITLE The sequence of A. thaliana F17123
JOURNAL Unpublished (1999)
REFERENCE 3 (bases 1 to 134784)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63108, USA
e-mail: twilson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by
M. Iodhi, Cold Spring Harbor Laboratories, and fingerprinted
by M. Marra, Mashu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

Actual start of this clone is at base position 1 of F17123; actual
end is at 134784 of F17123.
Location/Qualifiers
1. 134784
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/clone="F17123"
/chromosome="IV"

BASE COUNT 43094 a 23410 c 23696 g 44584 t
ORIGIN
/map="unknown"
Query Match 2.9%; Score 191.2; DB 8; Length 134784;
Best Local Similarity 56.7%; Pred. No. 9.5e-24;
Matches 430; Conservative 0; Mismatches 273; Indels 55; Gaps 2;
QY 5812 atcagaacctccccctgattgccaactagaggaagattgcaatagacctcttaagg 5871
Db 46525 AGCTTAACCGGATTTGACTGGAACATGAGAAAGATACCAATTCATTCGACGAGAGG 46584
QY 5872 cctatcctaattgcatgacattgacccaagattaccatcgcatgtaaaagctgc 5931
Db 46585 TTTGTGTATCTACTATGAGCATGTGATCCCAAGATCATTCATAGAGATGAAGGAGC 46644
QY 5932 aaatatattatgacgaagaattgagcctgtgtagtgatttgggttaagtagct 5991
Db 46645 TAATATTCCTTAGACGAGTGTCTTGAAGCTGTGTGGTGACTTTGGACTCGCAAGCT 46704
QY 5992 catgattacaagatgaccatgtaagctgtaaggggtaccattgggcacatagc 6051
Db 46705 CCTTAACCATGCGGATTCATGTCACACATGCGGTCCGTGTCAGGTTGACCATTTGC 46764
QY 6052 tcccgagtaacctcctgactggaagatcaccagagaagacccatgtccttggtatggat 6111
Db 46765 ACTGTAATATCTCTCCACTGTCAGTCTTCGAGAAACGATGTGTTGGGTTGGGAT 46824
QY 6112 aatgctcctagagctcattactgacagagagctttgattcttcctgccttcgcaaga 6171
Db 46825 ACTATTTGCTCGAGCTCATACACCGGACTGAGACTCTTACGTTGTGAACCGTTACCA 46884
QY 6172 tgatgatgtatg--tgttgatggtgtagtgcgcgggtgttccttggttaattat 6229
Db 46885 GAAAGAGCTATGCTTGAATGGTGAAGAAATGACACTGAGATACACTTTCTGATTAC 46944
QY 6230 ttcaacataatgcttactacttctgtgtgccccttgtttattctcgcctgatt 6289
Db 46945 TTCATAGATCAACCAATTAATTAATTTCTCCGATGCTGTGA-- 46990
QY 6290 tgattcctatgcatgttatgcatattgacccgcttgtaagtccttttaagttaaagcc 6349
Db 46990 -----ACATTTGGTTTGGTGAAGAAAT 47011
QY 6350 ttttgaagagaagaaagtgtgagatgctgctgacccctgacccaattacattg 6409
Db 47012 TACATGAAGATGAAGAAAGTGAAGGAACTATTGATCGAAGACTCGGAACCTACACGTA 47071
QY 6410 acacagaagtgtgacagcttattcaagtagcatctctgtaacccaaggttcgccaatg 6469
Db 47072 AGATTGAAGTTGAGAGATGTGCAAGTGGCTTTCTATGACACACAAATATTCGACGCTC 47131
QY 6470 agcggcctaagatgtcagaggtgagtcgcgaatgcttgaaggtgagtgcccttcagaagagt 6529
Db 47132 ATCGTCTTAATTAATGCTGAAGTTGTTTGATGCTTGAAGGCGAGATTACCGAGAT 47191
QY 6530 gggacgagtgycacaagaagtgtgaatcatcatcaagac 6567
Db 47192 GGGCTGCTTCGATTAACATTCACATTTCTACCATGCC 47229

RESULT 14
LOCUS AB005234 89779 bp DNA PLN 20-NOV-1999
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MBK5,
complete sequence.
ACCESSION AB005234
VERSION AB005234.1 GI:2264306
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone: Mitsu1 pl
ORGANISM Arabidopsis thaliana

[illegible]

SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS	1 (bases 1 to 110211) Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D., Fujii,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E., Feldylym,T.V., Buell,C.R., Kethum,K.A., Lee,J.J., Ronald,C.M., Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., Vanden,S.E., Umayan,L., Tallon,L.J., Gill,J.E., Adams,M.D., Carter,A.J., Cressey,T.R., Goodman,H.M., Somerville,C.R., Coppenhaver,G.P., Peuss,D., Nierman,W.C., White,O., Eissen,J.A., Salzberg,S.L., Fraser,C.M. and Venter,J.C.
TITLE	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
JOURNAL	Nature 402 (6763), 761-768 (1999)
MEDLINE	20083487
REFERENCE	2 (bases 1 to 110211)
AUTHORS	Lin,X.
TITLE	Direct Submission
JOURNAL	Submitted (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT	On Dec 17, 1999 this sequence version replaced gi:3738313. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (http://www.tigr.org/tdb/at/at.html).
<p>Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (http://arthur.washnigornl.gov/pub/sgriall), Genefinder (Phil Green, http://genomic.stanford.edu/GENSICANW.html), and NeplantGene (http://www.cds.dtu.dk/services/NeplantGene/), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curation based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats were identified by RepeatMasker (Arian Smilt, http://ftp.genome.washington.edu/RM/RepeatMasker.html). Genes are numbered from the top to the bottom of the chromosome.</p>	
<p>We thank the CSHL/WashU/ARI consortium for sequencing BAC clones F5P23, F5J6, T17A5, and T13JL16, the ESA group for sequencing clones T3J3d, and Scott Jackson, Jintang Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalek, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.</p>	
<p>This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.</p>	
FEATURES	Address all correspondence to: atetigr.org .
SOURCE	Location/Qualifiers 1..110211 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="2" complement(1..2793) /note="Sequence from clone F27L4" complement(356..401)
misc.feature	
repeat_region	

Tue Jun 27 14:47:57 2000

us-09-180-798-1.rge

Page 21

Qy	6111	taatgcccacagagccatctactgagcaagagagccttgcattgcgcgccttgacgaag	6170
Db	53610	TACTTTTGCACAGCCATCACACAGAAATGAGAGCTCTCGAATTGTGGCAAGTCTGTTAGCC	53551
Qy	6171	atgcctgcctgttatgct	6186
Db	53550	AGAAAGAGACTATGCT	53535

Search completed: June 23, 2000, 23:30:33
Job time: 43188 sec

CC of the embryo sac. The apomictic seeds and embryos thus produced
CC can be developed into plant progeny. This is useful in plant
CC breeding programs. Controllable and reproducible apomixis provides
CC many advantages in plant improvement and cultivar development in
CC the case that sexual plants are available as crosses with the
CC apomictic plant. Apomixis provides for true-breeding, seed
CC propagated hybrids and could shorten and simplify the breeding
CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match 99.9%; Score 6686.2; DB 1; Length 6695;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6692; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tcgaatgacgaatcgccctaccccttgatttngaaatactagttgtatcttgatt 60
DB 1 TCAGATGACGAATCGCGCTACCTTGATTNGAAATCTAGGTGATGATCTTGATT 60
QY agtttttgatactcttgctgaattctcttagagatgcaacgctcttcaattat 120
DB agtttttgatactcttgctgaattctcttagagatgcaacgctcttcaattat 120
QY agtttttgatactcttgctgaattctcttagagatgcaacgctcttcaattat 120
DB agtttttgatactcttgctgaattctcttagagatgcaacgctcttcaattat 120
QY 121 gagcccttgatgacgacaaagatcagcatgtttgacagagtgagctaaagta 180
DB 121 GAGCCCTTGATGACAAAGATATCTAGCATGTTGATCAGSAGTGTAAAGTA 180
QY 121 gaggtgttgatgacacataatgtttgagccctatggtctcaatgaattgatt 240
DB 121 GAGGTGTGATGACACATAATGTGTTGAGCCCTATGCTATCAATGAGTTGAT 240
QY 181 ggcgtgttgatgacacataatgtttgagccctatggtctcaatgaattgatt 240
DB 181 GGC GTGTGATGACACATAATGTGTTGAGCCCTATGCTATCAATGAGTTGAT 240
QY 241 gcaagatatactgtatcatcagctgtagtgaaggtatataaagaagaacatc 300
DB 241 GCAAGATATAGCTGTATATATGATGATGAGGTATATAAAGAAATGAACAT 300
QY 301 ctctcgtatgacatcattctctctctgacatagtaagaagtttgcacatacgt 360
DB 301 CTCTCGTATGACATCTTCTCTCTCTCCATAGTTAAGCAGTGTGTGCACATGACGT 360
QY 361 gaaactgagatgctgctctctccatcagttgattgatactgataagtgctcaact 420
DB 361 GAAACTGAGATGCTGCTCTCTCCATAGTTAAGTGTGATGATGCTCAACTTC 420
QY 421 ttgctgacgctttctctcgaattttcccaagacaagaattctttagttaagtatg 480
DB 421 TTGCTGACGCTTTCTTCCGATTTTCCCAAGACAAGATCTTGTAGTTAATGTTATG 480
QY 481 ctctcgtgtgctgtgctacatttaagaactctctgtttttaaagtgaagaagaac 540
DB 481 CTCTGCTGTGCTGTGCTACATTGAGAACTTACTCTGTGTTTAAATGAGAAAGAAAC 540
QY 541 ctacactttttctgtgttcccttataatcaccctcttgagaggtttagaacttta 600
DB 541 CTACCTTTTCTGTGTCTTCTTATATGATACACCTGCTTGAGGCGCTTTAGACTTTA 600
QY 601 tcaactaaactatcaatgtttacagacaagctatacgttttatcccccgcgcg 660
DB 601 TCAACTAAACTATTCATGTTTACAGACAGCTATACGTTTATCCCGCCCGCGCG 660
QY 661 acctgtagcaaaagaagcgctgaagaactgattatccggttttatataattaca 720
DB 661 ACCTGNGACAAAGAAGCGCTGATGAACGATTTAATCCGTGTTTATTAATTAACA 720
QY 721 ttgatgctcatgagactaatcttctgtttaattcaatgataatataacacctcct 780
DB 721 TTGATGCTCATGAGACTAATCTTGTGTTAAATTCATGATATATATACCTTCCCT 780
QY 781 ctgtgtaggcagtgccctcgttttaattagcgtacttaattatcgtatgatactgta 840
DB 781 CTGTGATGTCAGTGCCTCTCTCTTAAATAGCTTAAATATATCTGATGATCTGTA 840

DB 781 CTGTGATGTCAGTGCCTCTCTCTTAAATAGCTTAAATATATCTGATGATCTGTA 840
QY 841 tgccttgacagatgacatcagattatccattgttgcctacaataaataaac 900
DB 841 TGCTTGACAGATGATGATCAATATACATTTGCTGCTCAAAAATAAAC 900
QY 901 tctattatgttattcttttttgtaaagaatcactaatgagccttggtgaagcg 960
DB 901 TCTATTATGTTATCTTTTGTGTAACAAGTAATGATGAGCGCTATGTAAGCGG 960
QY 961 atgcattacaactcactcagactagcttgcaagatcccaaatgctctgcagagctgg 1020
DB 961 ATGATTACCAACTTACAGACAGCTTGCAAGATCCCAACATGCTCGAGACTGG 1020
QY 1021 atccaaccttgtaaccttgacatggttcaatgtagacatgtaacaatgaacagtg 1080
DB 1021 ATCCAACTTGTAACCCCTTGCAATGCTTCACTGATGATGATGATGATGATG 1080
QY 1081 ttataagagtagatgacacttcccttattatttttttagcaagttacgaatttact 1140
DB 1081 TTATAGAGTAGTAGTACTCCCTTATTAATTTTGTAGCAAGTACCAATATTTACT 1140
QY 1141 aatgagcagatgctcctttaaatttctttaaatttcttagtaagcagacatca 1200
DB 1141 AATGACAGATGCTCTTAAATATTTTCTTAAATTTCTAGTAAAGCAGCATGTA 1200
QY 1201 tcttaagtatcctactgaatttaagacataacatctttttaaatactatagtg 1260
DB 1201 TCTTAAGTATCTCTACTGAAATTTAAGACATTAATCTTTTAAATAATCTTATGAGT 1260
QY 1261 gtttttcgacagcagacatatcttttctgtaattcagaacaccttctcccg 1320
DB 1261 GTTTTTCGACAGCGCACATATATCTTTTCTGATATCAGAACCTTTCTCCG 1320
QY 1321 acgataaataataaagataaactccttgaactaaattttaaatttcttctctt 1380
DB 1321 ACGATAAATAATAAGATTAATCTCTGAACTAATTTTAAATTTCTTTCTTTT 1380
QY 1381 tgtctcttgacagaagttcttatgctcttctgtgaagaatcattctatgataatt 1440
DB 1381 TGTTCTTGAGAGAAAGTTGTTATGCTTTGTGAAAGTCAATCTATGATTAATTT 1440
QY 1441 tggcaactcataaattatataatcattccatgtagttaagttaaagaagcttcc 1500
DB 1441 TGCAACTCATTAATTAATTAATTAATTCATTAATTAATTAATTAATTAATTA 1500
QY 1501 ttaattccaagatagagttcattttaaagtttggcatalccatgagttttgaaga 1560
DB 1501 TTAATCCAAGATAGAGTTCAATTTTAATGTTGCGCATCCATGAGTTTGAATAATG 1560
QY 1561 cagaatattgttgagttatatttacttaccactttagtggtcatgtagtattctg 1620
DB 1561 CAGAAATTTGTGAGTTAATTTTACTTACCAACTTTAATGAGGTATCAGATCTTG 1620
QY 1621 ggaatgcagactatctgacaaattgttccctcttgccaggttgaaatcttacaact 1680
DB 1621 GGAATGCAGACTATCTGCGCATTTGCTTCCCTTGCCAGTTGAAATTTACATACT 1680
QY 1681 tgaagaacatatacttgaatgcttagttttaaagaacaaatgcttcaabact 1740
DB 1681 TGAAGAACATATCTGGAATGCTTAACTTTTATACAGCAATGCTTCAATATCT 1740
QY 1741 gttaaagtgtagaaaagttgacttctagctcgaagtggttgagataatcatatga 1800
DB 1741 GTTAAAGTGTAAGAAAGTTGACTTCTAGCTTACACAGTGTGCGAATAATCTATA 1800
QY 1801 agcaactaaagcgtggcaatttttcttatattcacaatgtaattgtaattgta 1860
DB 1801 AGCACTTAAAGCGTGGCAATTTTGTATATTTTCAAAATATGTAATGTTACTA 1860
QY 1861 cttaatatgataaactgaatttaactcctcatgattggtcctcagtcgaatggtcc 1920
DB 1861 CTTAATATGATTAATGATTAATCTCTCATGATGATGCTCAATGATGCTCAATG 1920

OY	1921	agccacaatnaataactctgmgggtctggacaataataactctcttcttaaggtccagaa	1980
Db	1921	agtcacactmtatamaattggcgggttggaacaaatataaactctcttcttaaggtccagaa	1980
OY	1981	gagcactatcaaacctctgtctcagataacgtcaacagtgaggctacagaggtatcca	2040
Db	1981	gagcactatcaaacctctgtctcagataacgtcaacagtgaggctacagaggtatcca	2040
OY	2041	gtttgggaaggtttaatgagcacactattaccctgtctctttaaacyctcagagatgta	2100
Db	2041	gtttgggaaggtttaatgagcacactattaccctgtctctttaaacyctcagagatgta	2100
OY	2101	ttaagctcgtcatcattccagagtttaataatgaoccttcaagtctatgatgaatgata	2160
Db	2101	ttaagctcgtcatcattccagagtttaataatgaoccttcaagtctatgatgaatgata	2160
OY	2161	tgaagaatacatatcttaactgtctccatgscgtcttccaacatgctccataatctgtta	2220
Db	2161	tgaagaatatacatatcttaactgtctccatgscgtcttccaacatgctccataatctgtta	2220
OY	2221	tctttgtcaatccttaaaaaatggacatgaataaaatgtgagaagaagtagctctccaatcca	2280
Db	2221	tctttgtcaatccttaaaaaatggacatgaataaaatgtgagaagaagtagctctccaatcca	2280
OY	2281	tttcatctatccagagagaatacatataattttttaatacaatgttggccctagaagtt	2340
Db	2281	tttcatctatccagagagaatacatataattttttaatacaatgttggccctagaagtt	2340
OY	2341	tctcaagtaatggtcattttaattttccaccattagaacctgtgttcagatgaanaac	2400
Db	2341	tctcaagtaatggtcattttaattttccaccattagaacctgtgttcagatgaanaac	2400
OY	2401	ttagaactccacagaataatcttaagtaaaagatctttaagatctggtgtgaagaagt	2460
Db	2401	ttagaactccacagaataatcttaagtaaaagatctttaagatctggtgtgaagaagt	2460
OY	2461	tcatgtgtgtggccgtgccacagaaattiaaalcacatcccatgcatcagtgtgtgtgacga	2520
Db	2461	tcatgtgtgtggccgtgccacagaaattiaaalcacatcccatgcatcagtgtgtgtgacga	2520
OY	2521	ctactcaactgttaatcgaagaagaactcattgtgaaatgaatgactctttcttctgtctt	2580
Db	2521	ctactcaactgttaatcgaagaagaactcattgtgaaatgaatgactctttcttctgtctt	2580
OY	2581	catgttagcgcttaataaggtcctacgtcgtacatggtttttaaactgttatagttggt	2640
Db	2581	catgttagcgcttaataaggtcctacgtcgtacatggtttttaaactgttatagttggt	2640
OY	2641	actgacaaggtttaaagttctctctgtgttiacgaattagaatatataataataaacycttt	2700
Db	2641	actgacaaggtttaaagttctctctgtgttiacgaattagaatatataataataaacycttt	2700
OY	2701	aacttctctctggaagaagtgcttcttacccttttatatatatatagatactacaacct	2760
Db	2701	aacttctctctggaagaagtgcttcttacccttttatatatatatagatactacaacct	2760
OY	2761	gctggacaattatcatcttaacgaacttaacgaatatacagaagaacttgtatatagttcaagt	2820
Db	2761	gctggacaattatcatcttaacgaacttaacgaatatacagaagaacttgtatatagttcaagt	2820
OY	2821	agtggcctgtagaagaacaccttaagaagaacttaatcatgagaggttccaacctttaact	2880
Db	2821	agtggcctgtagaagaacaccttaagaagaacttaatcatgagaggttccaacctttaact	2880
OY	2881	tctcttttagagtttttccaagtttatggaanaattgtaacctcaagatcggtggttcttcc	2940
Db	2881	tctcttttagagtttttccaagtttatggaanaattgtaacctcaagatcggtggttcttcc	2940
OY	2941	ataaacttccaataagatccggttcttctgscgttttccaatgtaagatcggttgacagatg	3000
Db	2941	ataaacttccaataagatccggttcttctgscgttttccaatgtaagatcggttgacagatg	3000

QY	3001	atcaggcggttccttccaataatcaataatggtccacgttgatagcgctgtactattat	3060
Db	3001	ATTAGCGGTTCTTTCATTAATCAATATGTCGTCACTTTGAAGAGCGCTGTACTATTAT	3060
QY	3061	tgcacctgacactaaccttgatcccatcgtcatcttgatctgaatgctactaaccca	3120
Db	3061	TGCACCTTGCACTTAACCTTGATCCCATGTCATCTGAATTCATTAATCTTAACCGA	3120
QY	3121	gttgaacatggtttatcaatgctctttgaggtaacaatgtagcttaccctgtcttga	3180
Db	3121	GTTGACATGGTTATCATGTCTTTGAGGTAAACAATGACGTTACACCTGTCTCTGA	3180
QY	3181	tataggtttaaggtttccactcccatcgctcttcgttggtttttatcaagttcaaca	3240
Db	3181	TATAGGTTTAAGGTTCCACTCCCATCGCTCTTCGTGTTTATTTACAGTTCAACA	3240
QY	3241	ccctactagcactgtgcaccctcctagctctttgtccgcaaataglaagaagttcttcgca	3300
Db	3241	CCCTACTAGCACTGTTCACGCTCTAGCTTTGTGCGCAATAGTAAGAGTTTCTTTGCA	3300
QY	3301	taataatgtagatgatcatttaagaacaatggaatcaatatacgtttatgtgtttgtac	3360
Db	3301	TAAATGTGAGATCAATTTAAGAAATAGCAACAATTAATGTGTTATGTGTTGTAC	3360
QY	3361	tttggaaatcaatagttgtccttgacaactgtgtcgtttatcgttttcaaggcttgcaca	3420
Db	3361	TTTGGAAATTAATAGTTGCTGCAACATGTGTGCTTTATCTGTGTCAAGCTTTGCCAA	3420
QY	3421	ggaagcgagatagtaagaagtgggcattccaaagcgcttactttgaaggcgcgcgac	3480
Db	3421	GGAAAGCGCATAGTAAGAAGTGGGCATCCAAAGCGCTTATCTTGAAGGCGCGCGCAC	3480
QY	3481	gtttgaggatttcgggtgtcctatagaaggacattatcatatactggtttatattaga	3540
Db	3481	GTTTGAGATTCTGGGTGTCTATTAGAGACATTATCTAATTAATCTGATTATTATTGA	3540
QY	3541	atataaatcaactactaatacttcttgytaatggtttatataagaatcccaactcgtaac	3600
Db	3541	ATAATAATCAACTCTAATATTTTCTTTGAATGTTAATAGAAATCCCACTCGTAAC	3600
QY	3601	ttgaaaataccattgaaatatttgaaccatataattgtagtgtgcaggtttaatcca	3660
Db	3601	TTGAACAATACCATTGAAATATTTGAACCTAATTAATATGATGTGTCAGGTTAATATCA	3660
QY	3661	aacctattaatcttactttaaaaaataatctcatatgaaatcgtaaacagtaataatat	3720
Db	3661	AACCTATTAACTTACTTTAAAAAATTAATCTAATGAAATGTAACGATATAAATATAT	3720
QY	3721	taaatltaatgtagtgtgctctatataatagctgtaatgcttaatagacttccaaagcgctg	3780
Db	3721	TAAATTAATGATATGATGTGCTATATTAATTAATGCTAATATAGATCCCAAGAGCGCTG	3780
QY	3781	ccttaacgacctagagcttccagggaagttaacatgcttaactttgaacaaatctgggtctc	3840
Db	3781	CTCTTAAGCTCTAGGCGTCCAGGAGATTCCTATGCTTAATCTTGACAAATATGGGGTTC	3840
QY	3841	gtagacatglttgaggatccctcatcactgtaactccgttttgctgacccctcgttcat	3900
Db	3841	GTAGACATGTTGTGGGATCCCTATCACTAGTAATCCGTGTTTGCTGACCCCTGTTCAT	3900
QY	3901	tattattattgtagtgaattactaagtattataaaatatctttatttgaataaattaac	3960
Db	3901	TGATTTTCAATGATGATGATTAATTAATCTTTTAAATTAATCTTTATGCAAAATTAATAC	3960
QY	3961	tggagtttaaaatgacagggagctttacagcaataaataagtgagcaaatcttccatag	4020
Db	3961	TGGAGTTTAACAAATGACAGGAGCTTTACAGCAATTAATAGTGAGCAAAATCTCTAGTG	4020
QY	4021	accttgggaacctgacaatcttgtagcgcttgtagcctataacatgaaatagcttcttgac	4080
Db	4021	ACCTTGGGAACCTGACAAATTTGGTAGCTTGGAACCTTAATACATGAATACCTTCTGTGAC	4080
QY	4081	ctataccggaacatattagaagaagttcaagqctaaatcttctgtatgactcaaatctt	4140

```
Db 4081 CTATACCGGACACTTTGGAAAGCTTACAAAGCTTAAGATTCTGTATGACTTACAAATCTT 4140
Qy 4141 cactagtttttaacttaactgaatlgatlatcccttcaagtgatlgatataccaa 4200
Db 4141 CACTAGTTTTTAATCTATGCAATTTGATATATCTTCAAGTGATATATATACAAA 4200
Qy 4201 ttactgatatggcgtcccaaaactgcctctctgtgccaatccaatgtaactgact 4260
Db 4201 TTAGTGTGTTCAATTTTAATACTAATAATGTTCAACGAGATTTATCAAAACATCGCTAT 4260
Qy 4261 aataaactcctcgaagtcctgtaaglatccgaaccttccgaagtgatgtgtg 4320
Db 4261 AATTTTACACTCTTCAAGTCTTAAGTATTCGACCTTCCGAGATGTTGTGTG 4320
Qy 4321 tgaatgttcaactttaactaaataatgtcacaaggaatttcaacaactgcgcat 4380
Db 4321 TGGATGTTTCAATTTTAATACTAATAATGTTCAACGAGATTTATCAAAACATCGCTAT 4380
Qy 4381 cagagaccgaatccggataatggtcatttcttggtaaccataagtttaagtcta 4440
Db 4381 CAGAGACAGTACCGGATTAATGGCTATTGTTGTTTACACCTATCAGGTTTAATGCTA 4440
Qy 4441 gtaataccttaactaataatggttcttactcactgcgaagcatgataatatttc 4500
Db 4441 GTAATATCTTAATATATATATGTTCTTACTTCTACTGCGAAAGCATATATATTTTTT 4500
Qy 4501 tctccctcaataatacaactctgcgaagtttggcaaatattgaattatgtgacctg 4560
Db 4501 TCTCTCTCATATATATATATCACTTTCGAGTTTGGCAATTAATTTATGTTGAGACTG 4560
Qy 4561 taaactggagagccctgcctgcgagatctccccaatttctcccaaccctccctccac 4620
Db 4561 TAACTGGAGAGCCCTGCCCCGTGATCTCCCCATTTTCTCCACCCTCCGTTCAATCCAC 4620
Qy 4621 catcaacagtaacagcctcccaagtgatattgaatttataataatccgtaataattta 4680
Db 4621 CATCAACAGTACAGCCTCAAGCTGATTTAGTTTATTAATTAATCCGATTAATTAATTTA 4680
Qy 4681 tgaactgttaaaaaatgggtgttaattcaaccagttgcgaataaagtatttccctcttc 4740
Db 4681 TGACTGTAAAAATTTGGTGTATTTTCAACAGTTCCGATAAAGTATTTCTCTTCTCTC 4740
Qy 4741 ttctattattatgaagacaanaatggtcccaactggaactatgtctggggagtagctgc 4800
Db 4741 TTCTATTATATGAGAGACAAATATGTCCTCACTGAGACTTTCTGCGGAGTAGTCTGC 4800
Qy 4801 tgggtgcttacttgctgtctgcacactgcacatgtgcatttgcagtgtgagcgaagaana 4860
Db 4801 TGGTCTGCTTACTGTTGCTGTCACCTGCAATGGCATTTGCAATGTCGCGAGAAAAA 4860
Qy 4861 acccggaanaacatttcttggatgtgccaggttagtccgttaataagatatctatgagc 4920
Db 4861 ACCGGGAACAAATTTCTTGTGATGTGCCAGGTAGTCCGTGAAATAGATATCTATTGAGC 4920
Qy 4921 gcttaactgtcgtgagacttcttcaactgcatagtttaactcactcagtgagaagacc 4980
Db 4921 GCTTACTGTGTGAGACTTGTGTTTCACTGTCATTAGTTAATCACTGAGTGAAGGACCC 4980
Qy 4981 agaagtgacacttggttcaactgaagaggttctctcgcgagaattgcgaagtcgaacgga 5040
Db 4981 AGAAGTGACACTTGTGTCACAGAGAGGTTTCTCTCGAANAATTCGCAAGTCGCAAGGGA 5040
Qy 5041 taacttgatcacactccttggaagaagtggtgatttgtaagtgataaggaagcccttcg 5100
Db 5041 TACTTTTAGTACACTCTTGGAGAGGTGGATTTGGTAAGGTGATTAAGGACCCCTTGC 5100
Qy 5101 tgaatgctcacttgaagtagtaaaaggcttaagaagaagaacacgaagtgagact 5160
Db 5101 TGATGTGTCACACTGTGAGAGAGTTAAAGGCTTAAGAGAAAGAACACAGAGTGGCGAGCT 5160
Qy 5161 gcaagttcaacaagaagtggaaatgataagcaatgctgtgcacgaaactcttgcgctc 5220
Db 5161 gcaagttcaacaagaagtggaaatgataagcaatgctgtgcacgaaactcttgcgctc 5220
Db 5161 GCAGTTTCAACAGAGAGTGAAGTAATGATGATGAGCTGGCTGTCAGAAATTTCTGCGTCT 5220
Qy 5221 acgtgttctcgtcatgacacactaccagagcgctctctgtatataccataatgctaatgg 5280
Db 5221 ACGTGTGTTCTGTCATGACACCTACAGAGCGGCTCTCTGTATATCCATACAGAGGCTAATGG 5280
Qy 5281 aagtttcgtaatgtttaaagagatctcagtttaacaaactcaacttgccgaaggtt 5340
Db 5281 AAGTGTGGCTGATGTTTAAGAGGATCTCACTTACATTTCCATTAATCTCCGAAAGTT 5340
Qy 5341 tgttgatataaataalgaataaactccctcaactatglttaaggltgtataattctgag 5400
Db 5341 TGTGTTGATTAATAAATGAATATTAATCTCCCTACACTAGTATAGGAGTATATATTCGAG 5400
Qy 5401 cagatcttaattcccatgtgacaagataccagttatattgttttctcgttaattgatacg 5460
Db 5401 CAGATCTTAATTTCCCATTTGCAAGATACAGATCATTAATGTTTCTGTAATTAATACCG 5460
Qy 5461 gttataattctctcgtgatttggttatatgcaaggaatttcgagttcaataagttatcaa 5520
Db 5461 GTTATATTTCTTTCTTGTATTTGGTTATATGCAAGAGATTTGCAATCTAATATGATCA 5520
Qy 5521 actggatgctatglttattctcgtcaatltgaatctgtctcactgltgcgaataatata 5580
Db 5521 ACTGGATGCTATGTTTATTTCTGCAATGGAATCTTGTCTTCAATGTCGCAAAATATATAGA 5580
Qy 5581 tccaacttggaaatcacttataataactgtgtgaagtcgaactgttgacttcaatata 5640
Db 5581 TTTCAACTTGGAAATCATTTATATATATCTGTGTAAGTCACTGTGACTTTTCATCATTA 5640
Qy 5641 attagcttcaataatcagaatctgcctcgaagtgagtttaccgaacatactcaaaccttcc 5700
Db 5641 ATTAGCTTCATTAATACAGAAATCTGCTAGTACGTTTACGAGACATATCTAATACCTTTC 5700
Qy 5701 ttaagccctgataataatcgcgccacttaacttcaagtgtgcctcctcgaattc 5760
Db 5701 TTAATGCCCCGTGTATTAATATCGTCCCACTACTTATTCAGTTTGTCTGCTCTGAAATTT 5760
Qy 5761 ttgatctgaacttgatgtgtgttcttcaataatgaaggtgcacataagaacc 5820
Db 5761 TTGATGTGACATTTGAGATGCTTGTATTCATCAATATGAGGCTCAGCCATCGAAGCC 5820
Qy 5821 tccccctgtatggccaactcaaggagaaggaatgacactaagatcttcaaggggcctatcaa 5880
Db 5821 TCCCCCTGATTTGGCCAACTAGGAGAGAGATTCACATAGACTTCTTAGGGGCTATCTAA 5880
Qy 5881 attgaatgacactgtgatcccaagattacacatcgagatgtlaaagctgcaaatatact 5940
Db 5881 ATTGACATGACCAATTTGATCCCAAGATTAATCATGCGCATGTAAAAAGCTGCAAAATATATT 5940
Qy 5941 attgacgaagaatttgaagcggtgtgtagtgatttgggttaggttaggtcatgagata 6000
Db 5941 ATTGACGAAGAAATTTGAGGCTGTGTAGGTAATTTGGGTTAGCTTAGGCTCATGAGATTA 6000
Qy 6001 caaggaataccatgttacgactgcgtgaagggtlaacaaatgggcaacatagctcccgagta 6060
Db 6001 CAAGGATACCAATGTTACAGACTGCTGTAAGGGGTACCATGGGCAATAGCTCCGAGATA 6060
Qy 6061 cctctcgaactggaagaatgatacagaagaacggaatgtcttgggttataggataatgctcct 6120
Db 6061 CCTCTGACTGGAATAATCTATCAGAGAAACGATGCTTGTGTTTGGGATAAATGCTCTCT 6120
Qy 6121 aagagctcattactggaacgaaggcctttagatcttgcgccttggacacgaatgagatgt 6180
Db 6121 AAGAGCTCATTTACTGACACAGAGGCTTTTATATTTGCTGCTGCTGCAACGATGATGATGT 6180
Qy 6181 taatgttctgaattgggtlaatgtgtccgggtgttcccttgggtlaattatccacatata 6240
Db 6181 TATGTTGTTGGAATGGGATGATGTCCCGGGGTGCTCTTGGTATTAATTCACATATTA 6240
Qy 6241 gtccttaactacttgtgtgccccttgttttatttctcgtcgtatattgattctttagt 6300
Db 6241 GTGCTTACTACTTGTGTGTGGCCCTTGTGTTTATTTCTGCTGCTGATTTGATTTCTTAGT 6300
```


QY	6301	catgttatgcatatgaccgcgcgtttgcaatgctcttttaagtttaaaagccttttaagag	6360
Db	6301	CATGTATGCAATATGACCGCTTTGCCAAGTCCTTTAGGTTAAAGCCCTTTTGAACAG	6360
QY	6361	aaaagattgagatgcygcgcgtccgactccgcagcccaagaatcatcatgtagacagagaatt	6420
Db	6361	AAAAGTTGGAGAGCGCTGCGATCCTGACCTGCAGAACAAATTAATTGACACAGAAATT	6420
QY	6421	gagcagcttatcctaagtagcatcattcctctgtaccccaagggctgcgaatgagcgcctaag	6480
Db	6421	GAGCAGCTTATTCAGAGTAGCATTACTCTGACCCAGAGGGTTCGCCAATGGAGGGCCTAAG	6480
QY	6481	atgctcagaagtagtccgaagaagcttgaagtgatgtagcgccttcgcaagaagaatggagcagtg	6540
Db	6481	ATGCTCAGAAGTAGTCCGAAGGCTTTGAAGGGAAGGCGCTTGCAAGAAAGGGGACAGTGG	6540
QY	6541	caaaaattgagacatccatcaagaagcgtgaattagcccaatcagcaactctgtacatg	6600
Db	6541	CAAAAATTGAGACATCCATCAAGACGTGATTAAGCTCCACATCGAACTTCTTAAGG	6600
QY	6601	atcctagactcgcagacagaataactgcatgctttagaattatctgtgtccaagaataaacagca	6660
Db	6601	ATCCTAGACTCGACAGATACTTGCAATGCTTTGAATATCTGTGTCCAAGATTAACAGCA	6660
QY	6661	tataaatgtaatgaatatattttttatagtt 6695	
Db	6661	TATTAATGTAAATGAATAATATATTTTATTTATGTT 6695	
RESULT	2		
ID	V06571	V06571 standard; cDNA; 1814 BP.	
AC	V06571;		
DT	03-AUG-1998	(first entry)	
DE	Daucus carota SERK gene.		
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;		
KW	plant breeding; ss.		
OS	Daucus carota.		
FX	Key	Location/Qualifiers	
FT	CDS	94..1755	
FT		/*tag= a	
FT		/product= SERK protein	
PM	MO9743427-A1.		
PD	20-NOV-1997.		
PR	13-MAY-1997; E02443.		
PR	14-MAY-1996; GB-010044.		
PA	(NOVS) NOVARTIS AG.		
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ,		
DR	WPI; 98-086529/08.		
DR	P-PSDB: M47013.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 21; Pages 47-51; 123pp; English.		
CC	The sequence is that encoding SERK, a putative receptor kinase.		
CC	It may be used as part of a method of producing apomictic seeds		
CC	comprising: (a) transforming plant material with a nucleotide		
CC	sequence encoding a protein which in active form in a cell or		
CC	cell membrane renders the cell embryogenic; (b) regenerating		
CC	the transformed material into plants or carpel-containing		
CC	plant parts; and (c) expressing the sequence in the vicinity		
CC	of the embryo sac. The apomictic seeds and embryos thus produced		
CC	can be developed into plant progeny. This is useful in plant		
CC	breeding programs. Controllable and reproducible apomixis provides		
CC	many advantages in plant improvement and cultivar development in		
CC	the case that sexual plants are available as crosses with the		
CC	apomictic plant. Apomixis provides for true-breeding, seed		
CC	propagated hybrids and could shorten and simplify the breeding		
CC	process so that selfing and progeny testing to produce and/or		
CC	stabilise a desirable gene combination could be eliminated.		
CC	Apomixis allows plant breeders to develop cultivars with		
CC	specific stable traits for such characteristics as height,		
CC	seed and forage quality and maturity.		
SO	Sequence 1814 BP; 530 A; 3354 C; 415 G; 515 T;		

Query Match	8.2%	Score 550.4	DB 1	Length 1814
Best Local Similarity	80.6%	Pred. No. 9.1e-108		
Matches 722	Conservative	0	Mismatches 31	Indels 143
				Gaps 1
QY 5800	agaagcgtcaacccatccagaaaccccccgtatgtgscacaactagggagagatgtgcactagg	5859		
Db 1044	AGACGGTCACCCATCAGAAACCTCCCTTGATGTGGCCAACTGAGAAAGAGATGTGACCTAGG	1103		
QY 5860	atctctaggggccctactcaaatgtcatgacccattgtgatacccaagatltccactgcgca	5919		
Db 1104	ATCTGCTAGGGGGCTTTCTTATTTGGCATGACCATTGTGATCCCAAGATTAATCCATCGTGA	1163		
QY 5920	tgtaaaagctgcgaatatatttttgagcagaagaatttagagctttttagtgatttttg	5979		
Db 1164	TGTAAAAGCTGCAAAATATTTATGTGACCAAAATTTAGGCTGTGTAGGTGATTTTGG	1223		
QY 5980	gttagctaggctcatgtgatatcaagaagatacccatgtttagcgtctgtlaaggggtacac	6039		
Db 1224	GTTATGCTAGGCTCATGTGATTAACAAGATACCATGTATCAACTGCTGTAAAGGGGTACTT	1283		
QY 6040	tgggcaataagctcccgagttacctctctgacgtggaaatgtcatcagaagaagaccgatgtctt	6099		
Db 1284	GGGGTACATGTGCTCCGAGATACCTCTCGACTGGAAAGTCAACAAGAAAGACCGATGTCTT	1343		
QY 6100	tgtgttagggataatgtccctagagatcaattcctgagacagggccttttagctgtgtcg	6159		
Db 1344	TGTGTTATGGGATTAATGCTCTTAGAGCTCTTTACTGGAAGAGAGACTTTTATGTTGCTCG	1403		
QY 6160	ccttgcgaacgatgatgatgttatgttgttggattggtglatgtgltcccggtgttcctt	6219		
Db 1404	CCCTGCGAAGAGATGATGATGTATGTTGTGTGATTTG	1440		
QY 6220	ggttaattattcaacatattagtggtctactaacttgtgtgtgccccttgttttatttcc	6279		
Db 1440	-----	1440		
QY 6280	tgccgtatttgattctcttagcatgtatgatcatattgacccgttggaatgtcttttag	6339		
Db 1440	-----G 1440			
QY 6340	gttaaaagccttttgaagaagaaaaagtttgagatgtgtgtcgatccctgacctgcagaac	6399		
Db 1441	GTTAAAAAGCTTTTGGAAAGGAAAAAGTTGGAGATGCTGTGATCCGACTTGGAANAC	1500		
QY 6400	aattacatgacacagaagtttgagcagcttatcaagatagattacccttgcagccaggt	6459		
Db 1501	AATTACATTACACAGATTTGACAGCTTATTTCAAAGTACATTAATCTGTGACCAAGGT	1560		
QY 6460	tctgcgaattgagcggcctaaagatgtlcaagaggtgagtcgcaatgtcttgaaggtgagcctt	6519		
Db 1561	TGCGCAATGAGCGGGCCTTAAGATGTACAGAGTAAATCCGAATGCTTGAAGGTGATGGCCTT	1620		
QY 6520	gcagaaaaagtgaggacgagatggtgcacaaaagtgtgaatgtcatccatcaagacgtgaattgct	6579		
Db 1621	GCAAAAAAGTGGGACGAGTGGCAAAAAAGTAAAGATCAATCCATCAAGAGTGTGAATTTGCT	1680		
QY 6580	ccacatggaactctgaatgattcctcagagctgcagacagataactcttgacgtcttggatla	6639		
Db 1681	CCAATGGAACCTTCTGAATGAGATCCCTTAAGCTGAGACAGATTAATCTTGATGCTTTGAATTA	1740		
QY 6640	tctgttccaaagataaacaagcatataaatgtgaatgtgaatlaatatltttatgtgt	6695		
Db 1741	TCTGTGCTCAAGATTAACAGCATTAATAATGTGATGAATAATTTATTTTATTTGCTT	1796		
RESULT 3				
ID V06585				
AC V06585	standard; DNA; 4081 BP.			
DT 03-AUG-1998	(first entry)			
DE Arabidopsis thaliana SERK gene.				

QY	4984	agtgacacctgtgtcgaactgaagaggtttctctctgagagaattgcgaactgcgaacggatgac	5043
QY	5044	tttagtacc-----tcccttgaagaggtgattgtgttaagtgtaagggagccct	5097
Db	3321	AGTTCATCTGGACAGCTCAAGAGGTTTCTTGGCGGAGCTACAAAGGGGAGTGAAGC	3380
Db	3381	GTTTAGTACAGAACATTTGGGCGAGAGTGGGTTTGGGAAGCTCAAGGGGAGCCTT	3440
QY	5098	tgctgatggtcctacgttagcagttaaaggcttaagaagaacgaacacagtgagcga	5157
Db	3441	GGCAGACGGAACTCTGTTCCTCTCAAGAGACTGGAAGGAAGGCCAACTCCAGGTGGAGA	3500
QY	5158	gtctcagcttcaacacgaagttgaaatgtatgatacttggtgtgcatcgaatctcttcgcg	5217
Db	3501	GCTTCATTTCAACAAAGAACTAAGATGATTAAGTATGAGCAGTTCATCGAACTCTTGAG	3560
QY	5218	tctacgtgtttctgcatagcaccaccagcagcgcttcttgatatccatacatgtgctaa	5277
Db	3561	ATTTCAGGGGTTTCTGTTATGACCGACCGACAGATGCTTGTGTCTTCATATGGCCAA	3620
QY	5278	tggagagtggttcgctcagtgtaagaggtatctcagfttaacaattaccataacttgcagaa	5337
Db	3621	TGGAAAGTTCCTTCGTCGTCTCAGAGGTAAAACTTAACATTAACATCTTGTCCTC	3680
QY	5338	gtttgtttgattaaaaatgaatatatactccctacactatgttaagtggtataattct	5397
Db	3681	TCTCATTTACTTTGACGTGAAGTGTTTTTCATGTGTTCTTATNAGGTCATATATGTT	3740
QY	5398	g 5398	
Db	3741	G 3741	
RESULT	4		
ID	V06591		
AC	V06591;		
DT	03-AUG-1998 (first entry)		
DE	Arabidopsis thaliana SERK gene.		
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;		
OS	plant breeding; ds.		
KW	Arabidopsis thaliana.		
FT	Key	location/Qualifiers	
FT	CDS	195..2072	
FT		/*tag= a	
FT		/product= SERK protein	
FN	MO97A3427-A1.		
PD	20-NOV-1997.		
PF	13-MAY-1997: E02443.		
PR	14-MAY-1996: GB-010044.		
PA	(NOVS) NOVARTIS AG.		
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;		
DR	WPI; 98-086529/08.		
DR	P-PSDB: M47023.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 27: Pages 91-95: 13pp: English.		
CC	The sequence is that encoding SERK, a putative receptor kinase.		
CC	It may be used as part of a method of producing apomictic seeds		
CC	comprising: (a) transforming plant material with a nucleotide		
CC	sequence encoding a protein which in active form in a cell or		
CC	cell membrane renders the cell embryogenic; (b) regenerating		
CC	the transformed material into plants or carpel-containing		
CC	plant parts; and (c) expressing the sequence in the vicinity		
CC	of the embryo sac. The apomictic seeds and embryos thus produced		
CC	can be developed into plant progeny. This is useful in plant		
CC	breeding programs. Controllable and reproducible apomixis provides		
CC	many advantages in plant improvement and cultivar development in		
CC	the case that sexual plants are available as crosses with the		
CC	apomictic plant. Apomixis provides for true-breeding, seed		
CC	propagated hybrids and could shorten and simplify the breeding		
CC	process so that selfing and progeny testing to produce and/or		
CC	stabilise a desirable gene combination could be eliminated.		
CC	Apomixis allows plant breeders to develop cultivars with		

Query Match	Best Local Similarity	Score	DB	Length
Matches 523; Conservative 0; Mismatches 202; Indels 143; Gaps 1	60.3%; Pred. No. 8e-44;	248.8;	DB 1;	2089;
CC specific stable traits for such characteristics as height, seed and forage quality and maturity.				
50 Sequence 2085 BP; 568 A; 419 C; 503 G; 599 T;				
QY 5800 agagcgtcacgcacatcagaaaccccccctgattgagcgaactagagagagattgactaagg 5859				
DB 1361 AGAAGGCCACCGTCAACAACCTCCGTTGATGGCCAAACGGGAAGAAATCGCGCTAGG 1420				
QY 5860 atctcttagggcgctactcgaattgacgacattgtaaccgaattatccatcgga 5919				
DB 1421 CTCAGCTCGAGGTTTGCTTACCTACATGATACTGCGATCCGAAGATCATTCACCTGA 1480				
QY 5920 tgtlaaagctgcgaatatataattgagcgaagaatttgaagctgttgtagtgaaatttg 5979				
DB 1481 CGTAAAGCAGCAAAACATCCTCTTAGACGAAGAATTCGGACCGGTGTGGAGATTTCGG 1540				
QY 5980 gttagctcaggctcatgattatcaagagatcccatgtttcgcgtcgtgtaagggtaacat 6039				
DB 1541 GTTGTGGCAAACTTATGAGCTATTAATAAACCTCACTCACTGCAACACACAGTCCCTGGCACAT 1600				
QY 6040 tgggcacatgcctcccgagatccctccgcgtcgaagtcacacagaagaacagctgtctt 6099				
DB 1601 CGGTCACATGGCTCCAGAAATATCTCCAAACCGGAAATCTTCAGAGAAACCGACGTTT 1660				
QY 6100 tggttatgagataatgcctccctagagctcaatcctgagacagagggcttttgactctcg 6159				
DB 1661 CGGATACGGAATCATCTCTTAGAACTATATCAACAGACAAAGACCTTTCGATCTCGCTCG 1720				
QY 6160 ccttgagcaagatgatgatgttattgtgttgattgggtatggtccgggtgttccctt 6219				
DB 1721 GCTAGCTAAAGACAGACAGACCTCATGTATTACTTACAGC----- 1757				
QY 6220 ggttaattatccatattagtgcttactactcttggttgcccttbtthttatttcc 6279				
DB 1757 ----- 1757				
QY 6280 tgccgtgatttgattctcttgtatgtatgacatattgaccgcgtttgcaatgcttttag 6339				
DB 1757 -----G 1757				
QY 6340 gttaaaagccttttgaagaagaaagattgagatgctgctgcatactgcagctgcagaac 6399				
DB 1758 GTGAAGAGATTGTTGAAGGAGGAAGAGCTAAGAGATTGATAGTGATCCGATCTTCAACA 1817				
QY 6400 aattacattacacagagaagttgagcagcttattcaagtagacatctctgtaccagagt 6459				
DB 1818 AACTACAGAGAGAGAGAGAACTGGAACAAGTAACAAAGTGGCGCTTGCTATGCACCCAGGA 1877				
QY 6460 tcgcgaatgagcggcccttaagatgctcaagagtgtaagcgcgaatgcttgaagatgagcctt 6519				
DB 1878 TCACCATGTGAAGACCAAAAGATGCTGAAAGTTGTAAGGATGCTGGAAAGGATGGGCTT 1937				
QY 6520 gcaagaagaatgaggacgagttgagcgaagaagttgagatcattccatcaaagacgtlaaattagct 6579				
DB 1938 GCGAGAGAAATGGGACGAATGGCAAAAAGTTGAGATTGAGGAGAGAGATGATTGAGT 1997				
QY 6580 ccaatgaactcttgaattgatactcctagagctctgacagataacttgcattgcttttgaaat 6639				
DB 1998 CCAATCTTACTCTTGATGTGATTCTTGAATTCTTACTTCAATTTCACAGCGCTTGAGTTA 2057				
QY 6640 tctgttccaaagataaacaagacataaa 6667				
DB 2058 TCTGGTCCAAAGGTAAAAA----- 2085				
RESULT 5				
X23533 standard; DNA; 3842 BP.				

PT generating transgenic plants resistant to Xanthomonas
 PS Claim 14; Page 59-60; 67pp; English.
 CC This invention describes a method for conferring disease resistance in
 CC plants. The invention describes the use of novel genes and proteins
 CC belonging to the Oryza longistaminata and Oryza sativa receptor
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from
 CC cassava, maize and tomato are also described. The genes and proteins can
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in
 CC rice or tomato.
 CC Sequence 3293 BP; 937 A; 634 C; 701 G; 994 T;

Query Match 1.6%; Score 109.6; DB 1; Length 3293;
 Best Local Similarity 60.8%; Pred. No. 3e-14;
 Matches 197; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 5829 attgccaactagagagagattgacacagatctcttaaggagcctacccaattgcatg 5888
 DB 2406 AATGGGAACAGTGGCTAAATAATAGCCATAGAGCTGGCAAGGGCTTCTTATTGCAAC 2465
 QY 5889 accattgtagcccaagattatccatcgcatgtaaaaagctgcaatatatttgagc 5948
 DB 2466 AGGATGCTCCCTTGTGTAATCCACCGCATGTCAAGTCCACAAATATATGTTGACT 2325
 QY 5949 aagaattgaagctgttaggtgatttggttagctagcagctcatgattacaagata 6008
 DB 2526 CTGAACTTGAACCTATGTTGAGATTTGGATTAAGCCAACTACTTTCGTAACAAATGTA 2585
 QY 6009 cccatg---ttacgactcgtgtaagggttacacattgagcagatctccagagactct 6065
 DB 2586 CCTGTGAGTGCATGTCTGCAATTCGAGATCTTATGGCTACATGCTCCAGAAATATCAT 2645
 QY 6066 cgacggaagaagtcacacagagacagatgctgtgtaaggataatgctcctag 6125
 DB 2646 AACGCTGAATTTGATGAGAAACCGATGTGTATAGCTTGGAGTGTGTGAGAC 2705
 QY 6126 tcattactggacagagggcctttg 6149
 DB 2706 TTATAACAGCAGCAAGGCCAGTAG 2729

RESULT 8

RESULT 8
 ID T62124 standard; cDNA to mRNA; 3176 BP.
 AC T62124;
 DT 10-JUN-1997 (first entry)
 DE Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
 KW Plant; morphogenesis; regulation; short; stem; alteration;
 KW inflorescence; extraneous; gene; expression; transformation;
 KW increase; control; form; length; ds.
 OS Arabidopsis thaliana.
 FH Key Location/Qualifiers
 FT cds 51..2981
 FT /tag= a
 FT /note= "plant morphogenesis regulatory protein"
 PN J09056382-A.
 PD 04-MAR-1997.
 PE 24-AUG-1995; 216187.
 PR 24-AUG-1995; JP-216187.
 PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
 PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
 DR WPI: 97-206629/19.
 DR P-PSDB; W13408.
 PT DNA encoding plant morphogenesis regulatory protein - useful to
 PT yield plants with short stems or altered inflorescence
 PS Claim 1; Pages 6-10; 17pp; Japanese.
 CC The present sequence encodes an Arabidopsis thaliana plant
 CC morphogenesis regulatory protein (MRP), which can be used to yield
 CC a plant with, e.g. short stems or altered inflorescence. The MRP
 CC acts on a plant at a specific site for a specific period, and can
 CC therefore be used to regulate extraneous gene expression in a
 CC plant. The MRP's cDNA or genomic DNA can be used to transform a
 CC plant to increase its MRP expression, and therefore control the

CC form (particularly stem length) of the plant.
 SQ Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match 1.6%; Score 108.2; DB 1; Length 3176;
 Best Local Similarity 58.1%; Pred. No. 5.8e-14;
 Matches 191; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 5823 cccctgattgccaactagagagagattgacacagatctcttaaggagcctacccaattgcatg 5882
 DB 2272 CTCTGATTGGGACACACGGCTTATGATAGCATATGTCGACACAAAGCTTATGCTATAC 2331
 QY 5883 tgcattgaccattgtagcccaagattatccatcgcatgtaaaaagctgcaatatatttgagc 5942
 DB 2332 TACACCAATGACTGTAGAGCTGTTGACAGATCTTACAGAGACGTAAGTGTCCAACTTCT 2391
 QY 5943 tgaagcaagaatttgagcgtgtgtaaggatttggttagctagcagctcatgattaca 6002
 DB 2392 TGGACAAAGACTTATGAGGCTGTTGACAGATTTTGGAAATAGCAAAAGCTTGTGTGT 2451
 QY 6003 aggaataccatgtagcagctgctgtaagggttacacattgagcagacataagctccagagactc 6062
 DB 2452 CAAAGTCAATCACTTCAACTTACCTGATGGGACAGATAGGTTACATAGAACCCGAGTAG 2511
 QY 6063 tctgacggaagaagtcacacagagacagatgctgtgtaaggataatgctcctag 6122
 DB 2512 CTGCACTTTCACGGCTCAGTGAAGAAATCCGATGTGTATGGAATATGTCCTTCTTG 2571
 QY 6123 agctcattactggacagagggcctttgat 6151
 DB 2572 AGCTGTAAACCCGAAGAACCCGTTGAT 2600

RESULT 9

RESULT 9
 ID T62125 standard; DNA; 9295 BP.
 AC T62125;
 DT 10-JUN-1997 (first entry)
 DE Arabidopsis thaliana plant morphogenesis regulatory protein DNA.
 KW Plant; morphogenesis; regulation; short; stem; alteration;
 KW inflorescence; extraneous; gene; expression; transformation;
 KW increase; control; form; length; ds.
 OS Arabidopsis thaliana.
 FH Key Location/Qualifiers
 FT exon 1803..1881
 FT /tag= a
 FT 1882..2227
 FT /tag= b
 FT 2228..2366
 FT /tag= c
 FT 2367..2467
 FT /tag= d
 FT 2540..2643
 FT /tag= e
 FT 2468..2539
 FT /tag= f
 FT 2644..2715
 FT /tag= g
 FT 2716..2809
 FT /tag= h
 FT 2810..2878
 FT /tag= i
 FT 2879..2968
 FT /tag= j
 FT 2969..3040
 FT /tag= k
 FT 3041..3118
 FT /tag= l
 FT 3119..3190
 FT /tag= m
 FT 3191..3266
 FT /tag= n
 FT 3267..3338
 FT exon

```

FT      Intron      /tag- o
FT      3339. 3421
FT      /tag- p
FT      3422. 3493
FT      /tag- q
FT      3494. 3586
FT      /tag- r
FT      3587. 3655
FT      /tag- s
FT      3656. 3740
FT      /tag- t
FT      3741. 3812
FT      /tag- u
FT      3813. 3888
FT      /tag- v
FT      3889. 3960
FT      /tag- w
FT      3961. 4048
FT      /tag- x
FT      4049. 4120
FT      /tag- y
FT      4121. 4209
FT      /tag- z
FT      4210. 4281
FT      /tag- aa
FT      4282. 4349
FT      /tag- ab
FT      4350. 4421
FT      /tag- ac
FT      4422. 4508
FT      /tag- ad
FT      4509. 4580
FT      /tag- ae
FT      4581. 4706
FT      /tag- af
FT      4707. 4778
FT      /tag- ag
FT      4779. 4860
FT      /tag- ah
FT      4861. 4932
FT      /tag- ai
FT      4933. 5018
FT      /tag- aj
FT      5019. 5090
FT      /tag- ak
FT      5091. 5176
FT      /tag- al
FT      5177. 5248
FT      /tag- am
FT      5249. 5412
FT      /tag- an
FT      5413. 5481
FT      /tag- ao
FT      5482. 5576
FT      /tag- ap
FT      5577. 5648
FT      /tag- aq
FT      5649. 5726
FT      /tag- ar
FT      5727. 5800
FT      /tag- as
FT      5801. 5882
FT      /tag- at
FT      5883. 6011
FT      /tag- au
FT      6012. 6096
FT      /tag- av
FT      6097. 6182
FT      /tag- aw
FT      6183. 6268
FT      /tag- ax
FT      6269. 6354
FT      /tag- ay

```

```

FT      Intron      6891. 6974
FT      /tag- az
FT      exon        6975. 7328
FT      /tag- ba
FT      J09056382-A.
PD      04-MAR-1997.
PF      24-AUG-1995; 216187.
PR      24-AUG-1995; JP-216187.
PA      (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKUTSUO.
PA      (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
DR      WPI; 97-206629/19.
PT      DNA encoding plant morphogenesis regulatory protein - useful to
PT      yield plants with short stems or altered inflorescence
PS      Claim 6; Pages 12-15; 17pp; Japanese.
CC      The present sequence encodes an Arabidopsis thaliana plant
CC      morphogenesis regulatory protein (MRP), which can be used to yield
CC      a plant with, e.g. short stems or altered inflorescence. The MRP
CC      acts on a plant at a specific site for a specific period, and can
CC      therefore be used to regulate extraneous gene expression in a
CC      plant. The MRP's cDNA or genomic DNA can be used to transform a
CC      plant to increase its MRP expression, and therefore control the
CC      form (particularly stem length) of the plant.
SQ      Sequence 9295 bp; 2873 A; 1581 C; 1602 G; 3239 T;

```

```

Query Match      1.6%; Score 107.2; DB 1; Length 9295;
Best Local Similarity 55.7%; Pred. No. 1.3e-13;
Matches 205; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

```

```

QY      5823 cccctgattggccaactaggagagagattgcactagatctcttagggcctatctaatt 5882
DB      6538 CTCCTGATTGGGACACACGCGCTTAAGATGATGCTGCAGCACCAAGTTTACCTATC 6597
QY      5883 tgcattgacattgtatcccaagattatccatcgcatgtaaaagctgcaatatattat 5942
DB      6598 TACACCATGACTGTAGTCCAGAGATCATTCACAGAGAGTGAAGTCTCCACATTTCTT 6657
QY      5943 tggacgaagaattttaggctgtttagttagttagttagttagttagttagttagttagt 6002
DB      6658 TGGACAAAGACTTAGAGGCTGTTGACAGATTGGAATGCGAAAGCTTGCTGTGTGT 6717
QY      6003 aggatcccatgttaccgactgtctgaagggtaccattgggacactagctccgagatcc 6062
DB      6718 CAAAGTCACATCTTCACTTCACTGATGAGGCGACGATGAGTACATCAACCCGAGATG 6777
QY      6063 tctcgactgaaagtcatacagaagccgactgtcttggttatgagtaagtctcctag 6122
DB      6778 CTCGCACCTTCACGCGCTCAGTGAATCCGATGCTACAGTATGGAATAGCTTCTTG 6837
QY      6123 agctcatctgacagagagctttagatcttgcgccttgcagagatgatgttga 6182
DB      6838 AGTTGTTAACCCGAAAGAACGCCGTTGATGAGCAATCACTCCACCATCTGTTGTT 6897
QY      6183 tctgtttg 6190
DB      6898 CTTTCTTG 6905

```

```

RESULT 10
X07356
ID      X07356 standard; DNA; 4104 BP.
AC      X07356;
DT      21-MAY-1999 (first entry)
DE      Arabidopsis steroid receptor Bin1 DNA.
KW      BIN1; steroid receptor; receptor kinase; transgenic plant;
KW      brassinosteroid; disease resistance; crop protection;
OS      Arabidopsis thaliana.
PH      key
FT      CDS      97. 3687
FT      PN      W09859039-A1.
FT      30-DEC-1998.

```

PF 24-JUN-1998; U13100.
 PR 24-JUN-1997; US-881706.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Chory J, Li J;
 DR WPI: 99-081275/07.
 PR P-PSDB; M97819.
 PT New receptor kinase Bin1 involved in brassinolide signalling -
 PT useful for promoting increased yield and disease resistance in
 PT plants and for modulating oocyte maturation
 PS Claim 8; Page 49-52; 72pp; English.
 CC This DNA sequence codes for a novel plant steroid receptor kinase,
 CC designated Bin1 (see M97819), which is involved in the pathway for
 CC the synthesis of the plant steroid hormone, brassinolide. 18 New
 CC Arabidopsis dwarf mutants were identified that lacked the ability
 CC to respond to brassinolide, and were named bin mutants. The bin1
 CC mutations were used to map the gene to a small interval on
 CC Arabidopsis chromosome 4. Bin1 was cloned by map-based cloning.
 CC The Bin1 polynucleotide was identified within this interval by
 CC sequencing the wild-type and mutant alleles of this nucleic acid.
 CC Overexpression of Bin1 in transgenic plants provides plants
 CC characterized as having enhanced disease resistance, increased
 CC plant yield or vegetative biomass and increased seed yield.
 CC Expression of Bin1 may also increase resistance to pesticides.
 CC Inhibition of Bin1, e.g. using antisense oligonucleotides (AON), is
 CC used to render plants male-sterile, and to reduce their stature or
 CC yield, e.g. for creating dwarf varieties. Since Bin1 homologues
 CC may be involved in regulation of the menstrual cycle and uterine
 CC function, Bin1, antibodies and AON may be useful as contraceptives,
 CC for improving success of in vitro fertilisation and to prevent
 CC premature labour. Transgenic animals are also provided, and are
 CC models for studying steroid-receptor interactions or can be used
 CC to screen for therapeutic agents.
 SQ Sequence 4104 BP; 1061 A; 870 C; 930 G; 1243 T;

Query Match 1.3%; Score 86.2; DB 1; Length 4104;
 Best Local Similarity 53.9%; Pred. No. 2.9e-09;
 Matches 200; Conservative 0; Mismatches 168; Indels 3; Gaps 1;

QY 5825 cctatgagcccaactagagagagatgacatgacatcttcaggggctatactaatg 5884
 DB 3028 CTTAAATGTCACACGCGGAGATTGCCATGATCAGATCAGAGGCTTTCCTT 3087
 QY 5885 catgacattgatacccaagattccatcgcatgtaaaagctgcaaatattattg 5944
 DB 3088 CACCAACACGCACTCGCATATCATCCACAGACATGAATCCAGTATGTGTGCTT 3147
 QY 5945 gacgaagaattgagagctgtgtagtgatttgggttagctagagctatgataaag 6004
 DB 3148 GATAGAAATTTGGAAGCTCGGTTTCAGATTTCGATGCGACGCTGTAGTGGATG 3207
 QY 6005 gataccatgttca---cgactgctgaagggtacacattgggacatagctccagtag 6061
 DB 3208 GATACGATTTAAGCGTCATACATTAAGTGTGATACCGGGTACGTCCTCCAGAGTAT 3267
 QY 6062 ctctcgactggaagatcagagaaacagatgcttctgttatgagataatgctccta 6121
 DB 3268 TACCAAGATTTCAGAGTGTTAACAAAAGAGACGTTTAAATGATACGTTGCTTACTC 3327
 QY 6122 gagctcattactgagacagagagctttagatctctgccttcgacagcatgatgtt 6181
 DB 3328 GAGCTACTTCACGGGTAAACGGCAACGATTCACCGGATTTTGAGATTAACAACCTGTG 3387
 QY 6182 atgtgtgtga 6192
 DB 3388 GGATGGGTGA 3398

RESULT 11
 X23531
 ID X23531 standard; cDNA: 3045 BP.
 AC X23531;
 DT 17-JUN-1999 (first entry)

DE Maize Xa21 gene DT4 cDNA fragment.
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice; DT4;
 KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
 OS Zea mays.
 PN W0909151-A2.
 PD 25-FEB-1999.
 PF 17-JUL-1998; U14841.
 PR 13-AUG-1997; US-910386.
 PA (REGC) UNIV CALIFORNIA.
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,
 PI Wang G;
 DR WPI: 99-204431/17.
 PT New RRR polynucleotides and nucleic acid constructs - used for
 PT generating transgenic plants resistant to Xanthomonas
 PS Claim 12; Page 58-59; 67pp; English.
 CC This invention describes a method for conferring disease resistance in
 CC plants. The invention describes the use of novel genes and proteins
 CC belonging to the Oryza longistaminata and Oryza sativa receptor
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from
 CC cassava, maize and tomato are also described. The genes and proteins can
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in
 CC rice or tomato.
 SQ Sequence 3045 BP; 717 A; 739 C; 744 G; 735 T;

Query Match 1.2%; Score 80.4; DB 1; Length 3045;
 Best Local Similarity 53.7%; Pred. No. 4.5e-08;
 Matches 190; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

QY 5825 cctgattggccaactagagagagatgacatgacatcttcaggggctatactaatg 5884
 DB 2254 CTTGATTGGGCAACAAGAAATGCAATGATGATGCTCGCAAGAGAGCTGCTCTC 2313
 QY 5885 catgacattgatacccaagattccatcgcatgtaaaagctgcaaatattattg 5944
 DB 2314 CACCAATGTTGCACTCCACACATACACGCGGACATGAATGTCAGCAAGTCTCTT 2373
 QY 5945 gacgaagaattgagagctgtgtagtgatttgggttagctagagctatgataaag 6004
 DB 2374 GAGCATATATCTCATGCTCCTACGATCGGATTTGGGAATGGCGGCTCGTAATGCTGT 2433
 QY 6005 gataccatgttca---gactgctgaagggtacacattgggacatagctccagtag 6061
 DB 2434 GACTCATCTTAACCGTGGACCAAGCTTTAGGAACACCTGTGTATGTGCTCCGAGTAC 2493
 QY 6062 ctctcgactggaagatcagagaaacagatgcttctgttatgagataatgctccta 6121
 DB 2494 TTCCAGTCGTTATTTTGACACACTAAGGCGACGCTACAGCTAAGGCGTGTCTCTG 2553
 QY 6122 gagctcattactgagacagagagctttagatctctgccttcgacagcatgat 6175
 DB 2554 GAGCTTCTCTCAAGGAAAAAACCAATCATCCGACTGATTCGCGCAATATAT 2607

RESULT 12
 V06587
 ID V06587 standard; cDNA to mRNA: 981 BP.
 AC V06587;
 DT 03-AUG-1998 (first entry)
 DE Arabidopsis thaliana SERK LRR homologous EST clone.
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
 KW plant breeding; launcine-rich repeat; ss.
 OS Arabidopsis thaliana.
 FH Key Location/Qualifiers
 FT CDS 104..760
 FT /tag= a
 FT /note= "shows high homology to SERK"
 PN W09743427-A1.
 PD 20-NOV-1997.
 PF 13-MAY-1997; E02443.
 PR 14-MAY-1996; GB-010044.
 PA (NOVS) NOVARTIS AG.
 DT De Vries SC, Hecht VRG, Schmidt EDL, Van Holst GJ;

DR WPI: 98-086529/08.
DR P-PSDB: W47019.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 75-77; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 1.1%; Score 74; DB 1; Length 981;
Best Local Similarity 71.0%; Pred. No. 7.1e-07;
Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 956 agggatgcatcattacacacttaagactagcttgcaagatcccaaatgtcttcagag 1015
DB 193 AGGAGATGCTCTTACGCTTCGCCGAGATTGACAGATCCAGACCATGCTCCAGAG 252
QY 1016 ctggagatcccaacctgtgacaccttgacatggtttcatgtgacatgtaacaatgaaa 1075
DB 253 CTGGGATCCACTCTGTATATCTGTACTGCTTCATGTCACCTGTATACCAAGACAA 312
QY 1076 cagtgataaagatgta 1093
DB 313 CCGGCTCACCTGCTGTGA 330

RESULT 13
V06588
ID V06586 standard; cDNA to mRNA; 1106 BP.
AC V06586;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 142..798
FT /tag= a
FT /note= "shows high homology to SERK"
PN W09743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: W47018.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 71-73; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;

Query Match 1.1%; Score 74; DB 1; Length 1106;
Best Local Similarity 71.0%; Pred. No. 7.4e-07;
Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 956 agggatgcatcattacacacttaagactagcttgcaagatcccaaatgtcttcagag 1015
DB 231 AGGAGATGCTCTTACGCTTCGCCGAGATTGACAGATCCAGACCATGCTCCAGAG 290
QY 1016 ctggagatcccaacctgtgacaccttgacatggtttcatgtgacatgtaacaatgaaa 1075
DB 291 CTGGGATCCACTCTGTATATCTGTACTGCTTCATGTCACCTGTATACCAAGACAA 350
QY 1076 cagtgataaagatgta 1093
DB 351 CCGGCTCACCTGCTGTGA 368

RESULT 14
V06588

ID V06588 standard; cDNA to mRNA; 788 BP.

AC V06588;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 2..664
FT /tag= a
FT /note= "shows high homology to SERK"

Query Match 1.1%; Score 70.8; DB 1; Length 788;
Best Local Similarity 69.6%; Pred. No. 3.2e-06;
Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 956 agggatgcatcattacacacttaagactagcttgcaagatcccaaatgtcttcagag 1015
DB 97 AGGAGATGCTCTTACGCTTCGCCGAGATTGACAGATCCGACCATGCTCCAGAG 156
QY 1016 ctggagatcccaacctgtgacaccttgacatggtttcatgtgacatgtaacaatgaaa 1075
DB 157 CTGGGATCCACTCTGTATATCTGTACTGCTTCATGTCACCTGTATACCAAGACAA 216
QY 1076 cagtgataaagatgta 1093
DB 217 CCGGCTCACCTGCTGTGA 234

RESULT 15
V06590
ID V06590 standard; cDNA to mRNA; 1063 BP.

AC V06590;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 106..762
FT /tag= a
FT /note= "shows high homology to SERK"

PN W09743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: W47022.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 86-88; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;

Query Match 1.1%; Score 70.8; DB 1; Length 1063;
Best Local Similarity 69.6%; Pred. No. 3.5e-06;

	Matches	96;	Conservative	0;	Mismatches	42;	Indels	0;	Gaps	0;
QY	956	aggcgatgacattacaacttaacgaactagcttgaagaatcccaaatgtcctgcagag	1015							
DB	195	AGGAGATGCTCTTTACGCTTCGCCGAGTTAAACAGATCCGACCATGTTCTCCAGAG	254							
QY	1016	ctgggataccaacctgtgaaccttgacatgattcatgtgacatgttaacaatgaana	1075							
DB	255	CTGGGATCCAACTCTGTATCTGTATCTGTACCTGTGTTCCATGTCACTGTAAACAAAGACA	314							
QY	1076	cagtgattataagagctga	1093							
DB	315	CCGCGTCACCTCGGTGGA	332							

Search completed: June 23, 2000, 22:48:42
 Job time: 40318 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic acid - nucleic search, using sw model

Run on: June 23, 2000, 11:34:14 ; Search time 274.21 Seconds
(without alignments) 3173.662 Million cell updates/sec

Title: US-09-180-798-1

Perfect score: 6695
Sequence: 1 tctagatgacgaatcgcgc.....atataatattttatgtt 6695

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents-NA:*

- 1: /cgn2_6/prodata/2/lna/5A.COMB.seq:**
- 2: /cgn2_6/prodata/2/lna/5B.COMB.seq:**
- 3: /cgn2_6/prodata/2/lna/5C.COMB.seq:**
- 4: /cgn2_6/prodata/2/lna/5D.COMB.seq:**
- 5: /cgn2_6/prodata/2/lna/6.COMB.seq:**
- 6: /cgn2_6/prodata/2/lna/FCUS.COMB.seq:**
- 7: /cgn2_6/prodata/2/lna/Backfile1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109.6	1.6	1554	4 US-08-587-680A-24	Sequence 24, Appl
2	61.6	0.9	5733	3 US-08-473-553A-1	Sequence 1, Appl
3	61.2	0.9	2571	1 US-07-717-331F-9	Sequence 9, Appl
4	61.2	0.9	2833	1 US-07-717-331F-1	Sequence 1, Appl
5	61	0.9	2749	1 US-07-717-331F-4	Sequence 4, Appl
6	60.6	0.9	2749	2 US-08-265-628-1	Sequence 14, Appl
7	59.4	0.8	7218	1 US-08-232-463-14	Sequence 13, Appl
8	55.6	0.8	19134	4 US-08-487-826B-13	Sequence 83, Appl
9	54	0.8	4517	6 PCT-US93-06251-83	Sequence 2, Appl
10	53.8	0.8	966	1 US-08-447-185-2	Sequence 3, Appl
11	53.8	0.8	2443	1 US-08-447-185-3	Sequence 77, Appl
12	50.8	0.8	2647	6 PCT-US93-06251-77	Sequence 1, Appl
13	50.8	0.8	2890	1 US-07-928-464-1	Sequence 1, Appl
14	50.8	0.8	2890	6 PCT-US93-07347-1	Sequence 1, Appl
15	50.8	0.8	3033	1 US-08-003-311B-1	Sequence 1, Appl
16	50.8	0.8	3033	1 US-08-261-432-1	Sequence 1, Appl
17	50.4	0.8	1203	2 US-08-602-010A-7	Sequence 7, Appl
18	50.4	0.8	1203	2 US-08-680-726A-7	Sequence 1, Appl
19	50.4	0.8	5495	2 US-08-602-010A-1	Sequence 1, Appl
20	50.4	0.8	5495	2 US-08-602-010A-2	Sequence 2, Appl
21	50.4	0.8	5495	2 US-08-680-726A-1	Sequence 2, Appl
22	50.4	0.8	5495	2 US-08-680-726A-2	Sequence 51, Appl
23	50.4	0.8	10592	2 US-08-680-726A-51	Sequence 52, Appl
24	50.4	0.8	10592	2 US-08-680-726A-52	Sequence 40, Appl
25	47.6	0.7	1804	2 PCT-US93-06251-82	Sequence 82, Appl
26	47.6	0.7	1804	6 PCT-US93-06251-82	Sequence 1, Appl
27	47	0.7	1602	1 US-07-820-011A-1	Sequence 1, Appl

28	47	0.7	1602	6 PCT-US93-00445-1	Sequence 1, Appl
29	46.6	0.7	5852	1 US-07-867-106-2	Sequence 2, Appl
30	46.6	0.7	5852	1 US-07-867-106-2	Sequence 2, Appl
31	46.2	0.7	6768	2 US-08-107-755A-1	Sequence 1, Appl
32	46.2	0.7	8457	1 US-07-991-867B-1	Sequence 1, Appl
33	46.2	0.7	8457	4 US-08-544-332-1	Sequence 14, Appl
34	45.6	0.7	2447	4 US-09-014-969-14	Sequence 36, Appl
35	45.4	0.7	665	4 US-08-883-795A-36	Sequence 1, Appl
36	44.4	0.7	6243	4 US-09-056-075-1	Sequence 12, Appl
37	44	0.7	1574	5 US-09-173-581-12	Sequence 1, Appl
38	44	0.7	3590	1 US-08-587-889-1	Sequence 1, Appl
39	44	0.7	3590	6 PCT-US96-09193-1	Sequence 3, Appl
40	43.8	0.7	1611	1 US-07-820-011A-3	Sequence 220, App
41	43.8	0.7	1611	6 PCT-US93-00445-3	Sequence 62, Appl
42	43.4	0.6	3209	1 US-08-105-483-220	Sequence 62, Appl
43	43.4	0.6	3209	1 US-08-220-151-62	Sequence 90, Appl
44	43.4	0.6	3209	1 US-08-413-118-62	
45	43.4	0.6	3209	2 US-08-224-391-90	

ALIGNMENTS

RESULT 1
US-08-587-680A-24
: Sequence 24, Application US/08587680A
: Patent No. 5977434
: GENERAL INFORMATION:
: APPLICANT: Ronald, Pamela C.
: APPLICANT: Wang, Guo-Liang
: APPLICANT: Song, Wen-Yuang
: APPLICANT: Szabo, Veronique
: TITLE OF INVENTION: Procedures and Materials for Confering
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/587,680A
: FILING DATE: 17-JAN-1996
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/373,375
: FILING DATE: 17-JAN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/475,891
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/004,645
: FILING DATE: 29-SEP-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/567,375
: FILING DATE: 04-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34,774
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1554 base pairs

Query Match	0.9%	Score 61.2	DB 1	length 2571
Best Local Similarity	53.1%	Pred. No. 6,28	05	
Matches 154	Conservative	0	Mismatches 133	Indels 3
			Gaps	1
QY 5049	gtaccatcccttggaagaagtgatattgtgtaagggtatataaggagacgcctgtgtatgct	5108		
Db 1592	gtAACCAACCTCGGACAGAGGTGGTTTGGATTGTTTAAAGAGGAAGATTACTTACCGGA	1651		
QY 5109	cactctgtagcagcttaaaaggcttaagaagaacagacacccagtgctcgagctgcagtttc	5168		
Db 1652	AAGAAATTGCGATTAATAAAGGCTTATCAAAACGTCAGTTCAAGGAGCAGTATG--AGTTTA	1708		
QY 5169	aaacagaagctgggaatgattagcatgctgtgtcatcogaatccttcgcttcaagtggtc	5228		
Db 1709	TGAATGAGGGTGACACTTAATTGGCGAGGCTTCACACATATAAACCTTGTTCAAGTTTGGCT	1768		
QY 5229	tctcatgacacactaccagacgagctctctgtatataccatatacgtgtaagtaagttc	5288		
Db 1769	GTTCGATTGAAGGAGATGAGAAAGATGTTGATATATGAGATTTGGAAAATTTAAGCCTTG	1828		
QY 5289	cgtaatgtttaagaaggtatctcagttacaatatcacaataactgcgcagaag	5338		
Db 1829	ATTGTATACCTCTTGGTAAACCCGGAAGCTTAAAGCTTAATTTGGAGATGAG	1878		

```

1  RESULT      4
2  US-07-717-331F-1
3  ; Sequence 1, Application US/07717331F
4  ; Patent No. 5484905
5  ; GENERAL INFORMATION:
6  ; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
7  ; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
8  ; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
9  ; TITLE OF INVENTION: A Receptor Protein Kinase Gene
10 ; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
11 ; NUMBER OF SEQUENCES: 9
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: Yahwak & Associates
14 ; STREET: 25 Skytop Drive
15 ; CITY: Trumbull
16 ; STATE: Connecticut
17 ; COUNTRY: USA
18 ; ZIP: 06611
19 ;
20 ; COMPUTER READABLE FORM:
21 ; MEDIUM TYPE: Floppy Disk
22 ; COMPUTER: Macintosh
23 ; OPERATING SYSTEM: MS-DOS
24 ; SOFTWARE: Microsoft Word
25 ; CURRENT APPLICATION DATA:
26 ; APPLICATION NUMBER: US/07/717,331F
27 ; FILING DATE: June 19th 1991
28 ; CLASSIFICATION: 800
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: George M. Yahwak
31 ; REGISTRATION NUMBER: 26,824
32 ; TELECOMMUNICATION INFORMATION:
33 ; TELEPHONE: (203)268-1951
34 ; TELEFAX: (203)268-1951
35 ; INFORMATION FOR SEQ ID NO: 1:
36 ; SEQUENCE CHARACTERISTICS:
37 ; LENGTH: 2833 base pairs
38 ; TYPE: nucleic acid
39 ; STRANDEDNESS: single
40 ; TOPOLOGY: linear
41 ; MOLECULE TYPE: DNA
42 ;
43 US-07-717-331F-1

```

	Query Match	0.9%	Score 61.2	DB 1,	Length 2833;	
	Best Local Similarity	53.1%	Pred. No. 6.4e-05;			
	Matches 154; Conservative	0;	Mismatches 133;	Indels	3;	Gaps 1
QY	5049	gtacatccttggaagggcgattggtataaggtataaggagccttgcgtatgct	5108			

Db 1592 GTAAACAACCGGACAAAGGGTTTGGTATGTTTACAAAGGAAGATTACTTGACGGGA 1651

QY 5109 cactctgtgaagtagttaaaggcttaaaagaagcaacacccagtgctgaagctgcagtttc 5168

Db 1652 AAGAAATTTGAGTAAAAAAGGCTATCCAAACCGTCAGTTCAAAGGAGCTATG--AGTTTA 1708

QY 5169 aaacagaagtcggaatcgattgattgcatgctgctgcacgaatcttcctgcctacgtgct 5228

Db 1709 TGATATGAGTGAGACACTAATTGCGAGGCTTCACACATATAAAGCTTGTCAAGTCTTGCT 1768

QY 5229 tcctgcatagcacactaccgagcggtctcttgatcatcatcacaatacctgctcgaagtgttg 5288

Db 1769 GTTCCATTGAGAGAGATGACAAATCTTATATATGATGAGATTATGAAATTTAAACCTTG 1838

QY 5289 cgtatctttaagaagatcctcgtttaacaattacccatacttcgcacagaag 5338

Db 1839 ATTCTTATCTCTTTGGTAAAAACCCGAAGGCTTAAGCTTAATATGGAATGAG 1878

```

? RESULT 5
? US-07-717-331F-4
? Sequence 4, Application US/07717331F
? Patent No. 5484905
? GENERAL INFORMATION:
? APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
? APPLICANT: Stein
? TITLE OF INVENTION: A Receptor Protein Kinase Gene
? TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
? NUMBER OF SEQUENCES: 9
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Yahwak & Associates
? STREET: 25 Skytop Drive
? CITY: Trumbull
? STATE: Connecticut
? COUNTRY: USA
? ZIP: 06611
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy Disk
? COMPUTER: Macintosh
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: Microsoft Word
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/717, 331F
? FILING DATE: June 19th 1991
? CLASSIFICATION: 800
? ATTORNEY/AGENT INFORMATION:
? NAME: George M. Yahwak
? REGISTRATION NUMBER: 26,824
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (203)268-1951
? TELEFAX: (203)268-1951
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2749 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? US-07-717-331F-4

```

	Query Match	Best Local Similarity	Score 61;	DB 1;	Length 2749;
	Matches 164;	Conservative	51.7%;	Pred. No. 7e-05;	
				Mismatches 150;	Indels 3; Gaps 1
QY	5829	atggccaactagggagagatggcactagatgctctcaggggcctatcaatitgaty	5888		
DB	1886	ATTGGCAAAATGAGATTGTGATTTATTAAGGATTTGGCCAGGGCTTCTATCTTAC	1945		
QY	5889	accatctgatcccaagatatccalcgcagatgtaaaagctgcacaatatatttggagc	5948		
DB	1946	AAGATTACGGTTTAGATTATCTACAGAGGATTTAAAGCAGGCAAGTCTCTGTTATA	2005		

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match 0.9%; Score 59.4; DB 1; Length 7218;
Best Local Similarity 1.6%; Pred. No. 0.00023;
Matches 6; Conservative 229; Mismatches 141; Indels 0; Gaps 0;

QY 299 ttctctcgtagcattcttctctctgctatagtaacgagtttgcacacatgacg 358
DB 1067 yy 1126
QY 359 ttgaactgtagtctgtctctccatcctaagtttgatctacgtatagatgctcaact 418
DB 1127 yy 1186
QY 419 tcttcgtagccttctcttcgattttcccaagaagaagattcttagtaatagttat 478
DB 1187 yy 1246
QY 479 tgctcgggagctgtgctgatttaggaatcttactctgtttttaaagsgaagaa 538
DB 1247 yy 1306
QY 539 acctacttttctctgtctccctttagatcaccctgtgagcgcttagactt 598
DB 1307 yy 1366
QY 599 taccacaataatcatgtttaccagaacaagctatagcttcccccccccgcg 658
DB 1367 yy 1426
QY 659 ggaactgagacaaca 674
DB 1427 yyyyyyygtaccAA 1442

RESULT 8
US-08-487-826B-13/C
Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbie Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelisen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOC KEY NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
AMTI-SENSE: NO
US-08-487-826B-13

Query Match 0.8%; Score 55.6; DB 4; Length 19124;
Best Local Similarity 45.9%; Pred. No. 0.0023;
Matches 345; Conservative 0; Mismatches 394; Indels 13; Gaps 4;

QY 1114 ttcttagcaagttacgaataatttactcaattgagcagatgtctcttaataatttctt 1173
DB 16212 TTTTTCCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16153
QY 1174 taattcttagcctaagcagacatcattcttaagatctctctcgaatttaagacataat 1233
DB 16152 ATATGCTTTGTTAAATGATATGATATGATATGATATGATATGATATGATATGATAT 16093
QY 1234 acctttttttaaataatcattagagtgcttttccgacagcagacatatctttt 1293
DB 16092 TATATATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 16033
QY 1294 ctgtaattcagaacaacctctcccgacgataaataataaataaataaataaataaataa 1353
DB 16032 TGTGTAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15973
QY 1354 taatttttatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1413
DB 15972 GTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15918
QY 1414 tgaagaatcattctatgataatttttggcaaccataataattatataattcattc 1473
DB 15917 TTTCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15858
QY 1474 tagtataagttataaaaaagctcctactaattccaagataagagttca-ttttatagt 1532
DB 15857 TTTATTTTAATTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15798
QY 1533 ttggacatcagatgatttttgaaatgcaagaatttttgagtaatttttactacca 1592
DB 15797 TTATGTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15741
QY 1593 acttttagcgcagcagtgatctctggaaatgacagcattctcggtaactgttctc 1652
DB 15740 ATTATTTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15681
QY 1653 ctggcagcttgaaaaaatttacaatactctgtaagacacatacttggaatgttaagt 1712
DB 15680 TTTTAAATGTTT---TTTATTTTCTTTCTTTGTTTATTTTAAATTTTAAATTTTAAATTTTAAATTT 15625
QY 1713 ttatacagacaatgcttcaatctctgtaaaagtgtgaagaagttgaactttagct 1772
DB 15624 TTTTATTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15565
QY 1773 tcagcagttctcggaataatcattgaagacacataaagaagctgggcaattttttgta 1832
DB 15564 AAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15505
QY 1833 ttattcaaatattgttaattgttactactta 1864


```

CORRESPONDENCE ADDRESS:
ADDRESS: Michael L. Goldman
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,185
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/111,078
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman M., Michael L.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1000
TELEFAX: (716)-263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2443 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-447-185-3

Query Match          0.8%; Score 53.8; DB 1; Length 2443;
Best Local Similarity 53.9%; Pred. No. 0.0028;
Matches 159; Conservative 0; Mismatches 127; Indels 9; Gaps 2;

QY 5010 ttctctcgcgaattgcgaacgcgaactcttaagtaacca-----tccctggaa 5063
DB 149 TTCCCTTAGTAGATTGGAGAGACCACTAATATTGATCACAAGTTTAAATTGGAC 208
QY 5064 gagtggatttggtaaggtgtataaggagcccttgctgtatgctcaactgttgcaagttca 5123
DB 209 ATGGGTCTTTGGGAGAGGTTTACAGAGGCTGTTTGCCTGATGGAGCAAGGTGGCCCTGA 268
QY 5124 aaagccttaagaagaacgaacaccaggtggcgagctgcagtttcaaacagaagtgagaa 5183
DB 269 AAAGCGCTACACCTGAGTCCTCAGACAGGTATGAG---AGTTGGAACAGAAATTGAGA 325
QY 5184 tgattagcatggcgtgcatcgaatcttctgcgtctacgttggttctgcatacacctta 5243
DB 326 CTCTCTCATTTTGGAGAGATCGCATCTGCTTTCATTGATAGAGATCTCTGTGATGAAGAA 385
QY 5244 ccgaagcgctcttgatataccatcagatggtaaggaagtgtgtgcgtcatgttt 5298
DB 386 ATGAGATATCTTATTTATAAATACATGAGAAATGGGAACCTCAGAGACATT 440

RESULT 12
PCT-US93-06251-77
Sequence 77, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
```

```

ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-77

Query Match          0.8%; Score 50.8; DB 6; Length 2647;
Best Local Similarity 51.3%; Pred. No. 0.014;
Matches 118; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 5903 aagattatcatgcgcgaatgaagaagctgcgaatatatatatttgagaagaattgagact 5962
DB 1732 AATTATATTCATAGATAGATCTCGCATCAGCAACATTTCTAGTGGCAATGGACTCATATGC 1791
QY 5963 gctgtagtgatttgggttagctaggtcagtcagttacgaagataaccatgtttagact 6022
DB 1792 AAGATTCTACTACCTCGGATGCGCCGATTGATGAAGACAAATGATGACAGCAGACAA 1851
QY 6023 gcttaagggtgacatttggtggcacatagctccgagttactctgcacggaaagcatca 6082
DB 1852 GGTCAAGATTCCTCCATCAAGATGAGCGCCCGAGGCGACCCCTTACGGAGGTTCACA 1911
QY 6083 gagaagaccatgtcttggttatggataatgctcctagagctaatatc 6132
DB 1912 ATCAAGCTCAGCGTGTGCTTTTGGATCTTACTCAGACAGACTGTGTAC 1961

RESULT 13
US-07-928-464-1
Sequence 1, Application US/07928464
Patent No. 5367065
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R.
TITLE OF INVENTION: Constitutive Triple Response Gene and
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz and
ADDRESS: No. 5367065r1s
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/07/928,464
FILING DATE: 19920810
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Suzanne E.
REGISTRATION NUMBER: 32,279
REFERENCE/DOCKET NUMBER: UPN-1086
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2890 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 118..2583
US-07-928-464-1
```

Query Match 0.8%; Score 50.8; DB 1; Length 2890;

Best Local Similarity 52.8%; Pred. No. 0.014; Mismatches 117; Indels 3; Gaps 1;

Matches 134; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

```
QY 5879 aaattgcatgaccattgtgattcccaagattatccatcgcatgtaaaagctgcaatata 5938
DB 2104 AATTATCTTCACAAATGCAATGCTCCCAATTTGTCATAGAGATCTAATAATCTCCAAACTTA 2163
QY 5939 ttattgagcaagaattgagctgtgtgagtgatttgggttagctgagctcatgagt 5998
DB 2164 TTGCTTGACAAAATAATATACAGTCAAGCTTTGATTTGGTCTCCGGGATTGAAGGCC 2223
QY 5999 tacaagatatacccatgttaacgactgctgttaaggggtacattgagcacatagctcccgag 6058
DB 2224 AGCAGCTTT---CTTCTCTGGAAGTCAGAGCTGGAACCCCGAGTGGATGGCACCAGAA 2280
QY 6059 taccctcgcactggaagatcatcagaagaagaccgactgcttgggttaaggaatagctc 6118
DB 2281 GTCCTGCGAGATGAGCCGCTAATGAAGAAGTCAGATGTGACAGCTTCGGGGTCACTTGG 2340
QY 6119 ctagaagctcatc 6132
DB 2341 TGGGAGCTTGTCTAC 2354
```

RESULT 14
PCT-US93-07347-1

Sequence 1, Application PC/TUS9307347

GENERAL INFORMATION:

APPLICANT: Ecker, Joseph R.

APPLICANT: Kiebert, Joseph J.

TITLE OF INVENTION: Constitutive Triple Response Gene and

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz and

ADDRESSEE: Norris

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/07347

FILING DATE: 19930805

CLASSIFICATION:

```
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Suzanne E.
REGISTRATION NUMBER: 32,279
REFERENCE/DOCKET NUMBER: UPN-1086
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2890 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 118..2583
PCT-US93-07347-1
```

Query Match 0.8%; Score 50.8; DB 6; Length 2890;

Best Local Similarity 52.8%; Pred. No. 0.014; Mismatches 117; Indels 3; Gaps 1;

Matches 134; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

```
QY 5879 aaattgcatgaccattgtgattcccaagattatccatcgcatgtaaaagctgcaatata 5938
DB 2104 AATTATCTTCACAAATGCAATGCTCCCAATTTGTCATAGAGATCTAATAATCTCCAAACTTA 2163
QY 5939 ttattgagcaagaattgagctgtgtgagtgatttgggttagctgagctcatgagt 5998
DB 2164 TTGCTTGACAAAATAATATACAGTCAAGCTTTGATTTGGTCTCCGGGATTGAAGGCC 2223
QY 5999 tacaagatatacccatgttaacgactgctgttaaggggtacattgagcacatagctcccgag 6058
DB 2224 AGCAGCTTT---CTTCTCTGGAAGTCAGAGCTGGAACCCCGAGTGGATGGCACCAGAA 2280
QY 6059 taccctcgcactggaagatcatcagaagaagaccgactgcttgggttaaggaatagctc 6118
DB 2281 GTCCTGCGAGATGAGCCGCTAATGAAGAAGTCAGATGTGACAGCTTCGGGGTCACTTGG 2340
QY 6119 ctagaagctcatc 6132
DB 2341 TGGGAGCTTGTCTAC 2354
```

RESULT 15
US-08-003-311B-1

Sequence 1, Application US/08003311B

Patent No. 544166

GENERAL INFORMATION:

APPLICANT: Ecker, Joseph R.

APPLICANT: Kiebert, Joseph J.

TITLE OF INVENTION: Constitutive Triple Response Gene

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and

ADDRESSEE: No. 54416615

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/003,311B

FILING DATE: January 12, 1993

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/928,464

```

; FILING DATE: August 10, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lori Y. Beardell
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: DPN-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 118..2583
; US-08-003-311B-1

```

```

Query Match          0.8%; Score 50.8; DB 1; Length 3033;
Best Local Similarity 52.8%; Pred. No. 0.015;
Matches 134; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

QY 5879 aaatgcatgaccattgtatcccaagaattatccatcgcatgtaaaagctgcaaatata 5938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2104 AATTATCTTCACATGCAATCGCAATCTCCAAATTGTCATAGAGATCTAAATCTCCAACTTA 2163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 5939 ttattggacgaagaatttgaagctgtttaggtgatttttgggttagctagctcatgcat 5998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2164 TTGGTTGACAAAAAATATACAGTCAAGTGTGTGATTGTCCTCGCGATTGAAGGCC 2223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 5999 tacaaagatacccatgttacgactgctgtaagggttacattggcacatagctccgag 6058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2224 AGCACGTTT---CTTCCCTCGAAGTACGACGCTGGAAACCCCGAGTGGATGCGACACAGAA 2280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 6059 tacctctgcactggaagatcatcagaagaagaccgaatgtcttgggttatgggataatgctc 6118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2281 GTCCTCGAGATGAGCCGCTAATGAAAGTCAGATGTGTACAGCTTCGGGGTCACTCTTG 2340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 6119 ctgaagctcattac 6132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2341 TGGGAGCTTGCTAC 2354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: June 23, 2000, 22:41:28
 Job time: 40034 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 11:29:27 ; Search time 6198.48 seconds
(without alignments)
4377.901 Million cell updates/sec

Title: US-09-180-798-1
Perfect score: 6695
Sequence: 1 tctagatgacgaatcgcg.....aattaattttttatggtt 6695

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query

SUMMARIES

No.	Score	Match Length	DB	ID	Description
1	271	4 0	532	79	AM233982 sf32g05.y
2	245.2	3 7	656	38	AA738547 SBRUKS TA
3	244.8	3 7	1202	84	B09168
4	207.6	3 1	555	61	B09168 t2E10-t7 so
5	197	2 9	538	62	AI1728030 BNLGH1962
6	191.8	2 9	597	51	AI1900110 sc10a04.y
7	191.8	2 9	703	51	AI1727837 BNLGH1920
8	184	2 7	471	47	AI1729440 BNLGH1133
9	183.8	2 7	380	43	AI1866453 EST344774
10	182.6	2 7	475	80	AI155643 A087P60U
11	178.2	2 6	714	92	AA6329855 N201132e
12	175.4	2 6	299	20	A0969390 LERM53TF
13	173	2 6	464	84	D25047 R1CR2976A.R
14	172.2	2 6	991	38	B217171 t2E10RE TAM
15	170.6	2 5	1435	38	AA738544 SBRUK2 SO
16	162.2	2 5	555	62	AA738545 SBRUK3 SO
17	160	2 4	383	81	AI1896277 EST265720
18	145.8	2 2	655	46	AA4733465 sh67e01.y
19	143.4	2 1	621	92	AA738546 SBRUK4 SO
20	142	2 1	490	47	A0969361 LERM53TR
21	133	2 0	474	46	AI14686325 EST244646
22	126.2	1 9	608	79	AI146897 sa1e91.y
23	124	1 9	548	51	AM234133 sf34f03.y
24	123.2	1 8	647	51	AI1738857 BNLGH1118
25	120.8	1 8	496	64	AI173055 BNLGH11700
26	119.4	1 8	631	64	AI1996242 701550775
27	118.2	1 8	412	79	AM067518 660013E08
28	118.2	1 8	304	35	AM310334 sf34f03.x
29	117.6	1 8	524	64	C22490 C22490.Rice
30	114.8	1 7	449	35	AM033961 EST277623
31	112.2	1 7	385	49	C22560 C22560.Rice
32	109.4	1 6	496	64	A0056335 A0056335
33	108.4	1 6	572	80	AM092144 EST185240
34	106.8	1 6	335	82	AM284352 LG1_275.D
35	106.6	1 6	428	28	AL095036 Arabidops
36	106.6	1 6	438	70	AA080670 EST012_Su
37	105.6	1 6	438	74	AW201545 sf04d10.y
38	105.4	1 6	314	79	D40033 R1C51756A.R
39	104.6	1 6	489	80	AM257240 EST305377
40	104	1 6	525	62	AM329840 N201116e
41	102	1 5	459	35	AI1897516 EST166959
42	100.2	1 5	432	63	C22359 C22359.Rice
43	99.6	1 5	592	79	AI1988078 sc96e12.y
44	98.8	1 5	608	79	AM234174 sf22b02.y
45	98.6	1 5	287	82	AW310202 sf32g05.x
					AL088599 Arabidops

ALIGNMENTS

RESULT	1
AM233982	
LOCUS	
DEFINITION	532 bp mRNA EST 13-DEC-1998
ACCESSION	AF33905.Y1
VERSION	GM.C1028
KEYWORDS	5', similar to TR:023921 023921 SOMATIC EMBRYOGENESIS
SOURCE	RECEPTOR-LIKE KINASE, mRNA sequence.
ORGANISM	AM233982
	AM233982.1 GI:6566309
	EST.
	soybean.
	Glycine max
	Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;
	euphyllophytes; Spermatophyta: Magnoliophyta; eudicotyledons: core
	eudicots: Rosidae: eurosids I; Fabales: Fabaceae; Papilionoideae;
	Glycine.
REFERENCE	1 (bases 1 to 532)
AUTHORS	Shoemaker, R., Keim, P., Vodka, L., Erpelding, D., Coryell, V., Rhaman, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rutter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M.,

TITLE McCann, R., Waterston, R. and Wilson, R.
JOURNAL Public Soybean EST Project
COMMENT Unpublished (1999)
 On Feb 18, 1999 this sequence version replaced g1:4297707.
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com
 High quality sequence stop: 403.

FEATURES
 source
 1..532
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1713"
 /clone_id="Gm-cl028"
 /tissue_type="roots of 'Superpod' plants"
 /lab_host="DH10B"
 /note="Vector: pluescript II XR; Site_1: EcoRI; Site_2:
 XhoI; The mRNA was isolated from roots of Glycine max
 'Superpod' plants generously donated by Dr. Gary Stacey.
 The seedlings were inoculated with Bradyrhizobium
 japonicus, strain USDA110 prior to harvest. StrataGene's
 cDNA synthesis kit (catalog number 200401) was used to
 synthesize the cDNA. First-strand synthesis was performed
 with 5'-methyl dCTP, hence the ligated cDNA was
 hemimethylated. A modification of StrataGene's
 first-strand synthesis primer was used. An 'anchor'
 nucleotide (V-A-C, or G) was added to the 3' end of the
 primer [GAGAGAGAGAGAGAGAGAGACTGCTCGAG(T)18V] to anchor
 the primer at the 5' end of the poly(A) tract. After
 second-strand synthesis, the cDNA ends were filled in with
 cloned Pfu DNA polymerase, ligated to EcoRI adapters and
 subsequently phosphorylated. The XhoI sites within the
 first-strand synthesis primer was then restricted by
 digestion with XhoI; all XhoI sites in the cDNA would be
 protected by their hemimethylated status. The cDNA
 constructs were size-fractionated with a 500bp cutoff,
 using GIBCOBRL Life Technologies' cDNA Size Fractionation
 column. The column eluent was then ligated into
 StrataGene's pluescript II XR predigested vector
 (pluescript II SK(+)) that has been digested with EcoRI
 and XhoI, and phosphorylated by StrataGene). Both the
 white and blue colonies appear to contain recombinant
 plasmids with cDNA inserts, based on size (n=25). This
 library was constructed by Dr. Paul Kelm and Dr. Virginia
 Corryell."

BASE COUNT	145 a	92 c	129 g	165 t	1 others
ORIGIN					
Query Match			4.0%;	Score 271;	DB 79;
Best Local Similarity			79.9%;	Pred. No. 1.2e-42;	Length 532;
Matches 319;			Conservative 0;	Mismatches 80;	Indels 0;
				Gaps 0;	
QY 5800	agagcgctgacgcctcgaacacctccccctgcatgctcaactcagaagagatgacctagg	5859			
Db 103	AGAAAGCTCAAGAAATCCCAACCCGCACTTGCTGGCCGCAACGAAAGCAAGCGATTCGATGGG	162			
QY 5860	atctctcagggggacctctcaaatgcatgacacattgcatcccaagaattatccatcgga	5919			
Db 163	ATCTGCAAGGGGGCTGCTTATTTCATGATGATTCATGGACCCATGAAGATTATTCACCGTA	222			
QY 5920	tgtcaaaagctgcacaatatattatgtgacagaagaatttgagggctgctgtgagtgatttgg	5979			
Db 223	TGTCAAGCAAGCTCAATATATTGTTGGACGAGAAATTTGAAGCAAGTTGTTGGAAATTTTGG	282			

QY	5980	gttaactgaggccatcggattatacaagatacccatgttaacgactgctgtgaagggtacat	6039
Db	283	TTTACCTAACTTAAGATATGAGTATAAAGATACATGATCTACTGCTGACGTGGAAACAT	342
QY	6040	tgggacatagctcccgagctacccctcgcactgagtggaagtcatacagaagaaccgagtctt	6099
Db	343	TGGACATATGACACACAGAAATACCTCTCAACGTGAAGTCTTCAGAGAAGACTGATGT	402
QY	6100	tgttataggataaagctccctcctagagctcattcctctgagagagggcttttgactctgc	6159
Db	403	TGGATATGCTGTGAAGCTTCTTGAACATATAACTGACAAAGGCGCTTTGATCATCTGC	462
QY	6160	ctctgacagcatgatatgtatgtttgttgatgggt	6198
Db	463	ACTTGCCAAATGATGATGATGATCTGCTGTTATTTGGG	501
RESULT	2		
LOCUS	AA738547		
DEFINITION	AA738547	656 bp	mRNA
VERSION	AA738547		
KEYWORDS	AA738547.1	GI:3421472	
SOURCE	EST.		
ORGANISM	Sorghum.		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllipytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.		
TITLE	1 (bases 1 to 656)		
JOURNAL	Annen,F., Chang,J.-L., Paterson,A.H. and Stockhaus,J.		
MEDLINE	Characterization of 14 different putative protein kinase cDNA clones of the C4 plant Sorghum bicolor		
COMMENT	Mol. Gen. Genet. 259 (1), 115-122 (1998)		
	On Nov 29, 1993 this sequence version replaced gi:636097.		
	Contact: Stockhaus J		
	Institut fuer Entwicklungsbiologie und Molekularbiologie der Pflanzen		
	Heinrich-Heine-Universitaet		
	Universitaetsstrasse 1, 40225 Duesseldorf, Germany		
	Tel: (49)-211-81-14953		
	Fax: (49)-211-81-14871		
	Email: stockhaeuni-duesseldorf.de		
FEATURES			
SOURCE	Insert Length: 1100	Std Error: 0.00.	
	Location/Qualifiers		
	1..656		
	/organism="Sorghum bicolor"		
	/strain="cv. Tx430"		
	/db_xref="taxon:4558"		
	/clone="SDRLK5"		
	/clone_11b="Sorghum bicolor cv. Tx430 leaf"		
	/dev_stage="green"		
	/note="Organ: leaf; Vector: lambda ZAP II; Site.1: Eco RI, Site.2: Xho I; leaf cDNA library from green leaves, unidirectionally cloned"		
BASE COUNT	161 a	145 c	184 g
ORIGIN			166 t
Query Match	3.7%	Score 245.2;	DB 38; Length 656;
Best Local Similarity	77.4%	Pred No.1,1e-37;	
Matches 311;	conservative	0; Mismatches 88;	Indels 3; Gaps 1
QY	5800	agagcgctacgcatcagaaacctcccctgattg	5856
Db	84	AGAGCGACGACATCTGACGCCACCTGTTAAGTGGAAACAACAAAGAAAGCGATTGCGCT	143
QY	5857	aggatcttcttaggggacctctcaaatgcatgacccattgtgatcccaagattatcatc	5916

Db	144	TGGAGTGTCAAGAGACATTCTTACTTCCAGATCAGATCGCATCCAAATATCATCCATACG	203
OY	5917	cgatgtaaagctgcgaatatattatggacgaagaatttgaagctgttgaagtattt	5916
Db	204	GGATCTCAAAAGCTGCAAAATATCTTTTCCAGAGAGACTTCAGGCAGTTGGGGATTT	263
OY	5977	tgggttaactgagctcattgattcaagaatgccatttcgactgtgtgaagsggtac	6036
Db	264	TGGGCTTCCCAAGCTTATGAGACTACAAAAGATACCATATGCATCAACTGTGTCCGTGGAC	323
OY	6037	cattggcacatagtctcccgagctactcttcgactggaagtcacatcagaagaaccgagt	6096
Db	324	AATTGGACACATTTGGCGCTGAGTACTTATCCACTGCGAAGTCTCTGAAAAGACTGATGT	383
OY	6097	cttgggttaagggaataatgctctcctagagctcattactggaagagggtttgactctgc	6156
Db	384	TTTTGGCTATGGGAGATCATGCTTCTTGAGCTTATTTACTGGCAGAGGCAATTGATCTTGC	443
OY	6157	tcgccttcgcaagctatgatgtattgtatttggattgggt	6198
Db	444	TGCTCTGCAATGATGACGAGTATATCTTCTTGACCTGGGT	485
RESULT	3		
LOCUS	B09168	1202 bp DNA	GSS 14-MAY-1997
DEFINITION	T2E10-T7 TAMU Arabidopsis thaliana genomic clone T2E10, genomic survey sequence.		
ACCESSION	B09168		
VERSION	B09168.1	GI:2090299	
KEYWORDS	GSS.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	1 (bases 1 to 1202)		
AUTHORS	Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.		
TITLE	BAC End Sequences at ANGC		
JOURNAL	Unpublished (1997)		
COMMENT	Other_GSS: T2E10-Sp6		
	Contact: Ecker J.		
	Arabidopsis Thaliana Genome Center		
	University of Pennsylvania		
	Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104		
	Tel: 215-898-9384		
	Fax: 215-898-8780		
	Email: jecker@atgenome.bio.upenn.edu		
	Seq primer: T7		
	Class: BAC ends		
	High quality sequence start: 76		
	High quality sequence stop: 704.		
FEATURES	Location/Qualifiers		
source	1..1202		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="T2E10"		
	/clone_11b="TAMU"		
	/sex="hermaphrodite"		
	/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"		
BASE COUNT	322 a 275 c 238 g 354 t	13 others	
ORIGIN			
Query Match	3.7%; Score 244.8; DB 84; Length 1202;		
Best Local Similarity	67.1%; Pred. No. 1.3e-37;		
Matches 397; Conservative	0; Mismatches 176; Indels 19; Gaps 3;		

QY	4728	ttctctctctctctctcttatattatagaagacaaatggtccactgagctatgctg	4787
Db	111	TTAACAATTTTGGTAAAAAATAAATTAAGTGGGTATGATCTCTCTGAGCCATTTGGG	170
QY	4788	gggagtagctgctgctgctgcttactcttcttgctgctgacctgcaatgcatgcatgctgct	4847
Db	171	GAGAGATGCTGCTGGTGGCTCTTACATTTTGCTGCTGCCCTCTCTTACCTTTTGCTTGGT	230
QY	4848	ggcggagagaagaaaccgcgagaaacattctcttgatgctgcaaggttagcttcctgtaataga	4907
Db	231	GGCCTAAAAAACAACCTCAAGATTTCTTTGATGTTCTCTGTAAGTCACATGATCTGCG	290
QY	4908	tatctatgtagagcgttac-----tgctgtagacttggtttcaactgcatc	4955
Db	291	AATATCCAGCTTTGTTTCAATTTTCAGAAATGGATTAATTAATTAATTTTTCAGT	350
QY	4956	agttaacctgactgtagaagagaccagaagtgcacctggtgctaactgtagaaggtttctc	5015
Db	351	TCCTGATGCAACCCGGAAGAGATCCTGAGGTTCACTTGGGGACGCTTAAGCGGTTCTCTC	410
QY	5016	tgcgagatgtcgaagtcgacgcgagatactttag-----taccatccttggaagagtg	5069
Db	411	TACGGGAATCTTAAGTANCACTGATAGCTTGACGAACAAAAACATTTTGGGCCGAGGTG	470
QY	5070	gatttgtagagtgatataaggaagcgcctctgctgtagtgctacttgttagcagttaaaagc	5129
Db	471	GGTTCGGAAAAAGCTCTCAAAAGGCCGCTCTTGATGGAACACTTTGTGAGCAAAACGCG	530
QY	5130	ttaaagaagaagcaaacaccagcgtggtgagctgacgttcaaaacagaagaatggaatgata	5189
Db	531	TTAAAGAAGCGCAACCCACGATGGCGAGCTCCAGTTTATAGACAGATGGAGATGATNA	590
QY	5190	gcattgctgtagcatgaaatctcttgctctacgttggtttctgcatg-acaactacagag	5248
Db	591	GCATGGCCGTTCAAGATATCTCCTCAGGCTACNTGGTTCTGTATGAACCTTATCCGAA	650
QY	5249	cgcgctctgtatatcatcatcatatgctgaatggaagtgttgctcatgtttaa	5300
Db	651	AGAAATGTTTGTATCTTCAATATGCTAAGGGAATAGTGGCTTCTGTTTGA	702

RESULT	4
LOCUS	AI728030
DEFINITION	AI728030 555 bp mRNA EST 11-JUN-1999
ACCESSION	BTGH196229 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic cottonogenesis receptor-like kinase [Daucus carota], mRNA sequence.
VERSION	AI728030
KEYWORDS	AI728030.1 GI:5048616
SOURCE	EST.
ORGANISM	upland cotton.
REFERENCE	Gossypium hirsutum
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
TITLE	1 (bases 1 to 555)
JOURNAL	Blewitt, M., Metz, F.C., Davy, D.F. and Burr, B.
COMMENT	ESTs from developing cotton fiber unpublished (1999)
	On Jun 5, 1998 this sequence version replaced gi:3187310.

```

FEATURES
  source      Location/Qualifiers
1. .555      1. .555
              /organism="Gossypium hirsutum"
              /cultivar="Acala Maxxa"

```

```
BASE COUNT      146 a      100 c      147 g      162 t
ORIGIN
```

Query Match	3.1%;	Score 207.6;	DB 51;	Length 555;
Best Local Similarity	62.2%;	Pred. No. 2.2e-30;		
Matches 432; Conservative	0;	Mismatches 119;	Indels 143;	Gaps 1

QY	5961	ctggttagtgatgatttgattgattgtagccatgcatgattcaagaatacccatgtaga	6020
Db	1	ctgtgtgttggtgactcttcttggttggttggcttaaccttattgaccttaccaggaatcccatgtaacta	60
QY	6021	ctctcgtlaagggtgtaaccattggcacacatagctcccgagtaacctctcgaactggaagtcac	6080
Db	61	ctgctgtatcgttgccacaaattggaacatatgtgcttccgagtaactctcgaactggaagtcac	120
QY	6081	cagagaagaacccgagctgctcttggttattggtgataatgctccctagagcttactgtaacga	6140
Db	121	cagaaaaaacctatcttcttcttgggttggtggtatcattcttcttggagcttcttgaactggaacgac	180
QY	6141	gggcttcttgatctctgcgcctctggacacgaatgcatgattatggttggattggtgatt	6200
Db	181	gggcttcttgatctctgcgcctctggacacgaatgcatgattatggttggattggtgatt	236
QY	6201	gtgtcccgagggtgtlcccttcttggttaattatccaatatagtgcttactactgtgtg	6260
Db	236	-----	236
QY	6261	ggccttgttttatttattccgtcgtctgatttgattcttctagtcattgcatgcatattgacct	6320
Db	236	-----	236
QY	6321	gcttctgcaatgctcttcttagatgtaaaagccttttgaaagagaanaaagtgtgagatgctgt	6380
Db	236	-----ggtcaaaagcactcttctgaaagagaanaaagctgtgagatgctgt	277
QY	6381	cgatctcgaacctgcgagaacaattacattggaacacagaagttagagcagctattccaagtgc	6440
Db	278	ttatctctgattctgcgaacccaattatgtatgtaaaactgaggttaaaagcgtttatccacaggttgc	337
QY	6441	attactctgtacccagggcttcgcacaatgtagagcgcctagaatggttcaagagttagtccgat	6500
Db	338	ttctgcttatgccaacaaagggttccccaattggaacacgagcgcgaacgaatgtcaaaagtgttaaaat	397
QY	6501	gcttcaaggtgtagtgcccttgcagaanaagtgggacgagtgcccaaaagttagaagtcatcca	6560
Db	398	gctggaagggtgattggttggccgaaagatgggatatgtaggcgaagaaagtgtgaagttctacg	457
QY	6561	tcaagaacgtagattatgctccacatcgagactctgaatgattccctgacatcgacagataa	6620
Db	458	gcagagaggtgtaacttgcccccctacatcccttaattctgatttgatgttgagactcaactacgaaa	517
QY	6621	cttgcatagcttttgataattatctgtccaagataa	6654
Db	518	ttctgcatgctgttgagttatccggtccaaaggta	551

RESULT	5
A1900110	
LOCUS	A1900110
DEFINITION	A1900110 538 bp mRNA EST 06-DEC-1999
ACCESSION	sclcd04.y1 Gm-cl012 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
VERSION	Gm-cl012-871 5' similar to TR:O23921 O23921 SOMATIC EMBRYOGENESIS
KEYWORDS	RECEPTOR-LIKE KINASE., , mRNA sequence.
SOURCE	A1900110 A1900110.1 GI:5606012
	soybean.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
ORGANISM	Glycine max			
	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eustosids I: Fabales: Fabaceae: Papilionoideae; Glycine.			
	1 (bases 1 to 538)			
	Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hiller, L., Kneub, T., Martin, J., Beck, C., Wyle, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, J., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ratter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.			
	Public Soybean EST Project			
	Unpublished (1999)			
	On Jun 5, 1998 this sequence version replaced g1:3169472.			
	Contact: Shoemaker R/Public Soybean EST Project			
	Public Soybean EST Project			
	Washington University School of Medicine			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: estewatson.wustl.edu			
	This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com			
	Seq primer: -40RP from glbco			
	High quality sequence stop: 411.			
FEATURES	Location/Qualifiers			
SOURCE	1..538			
	/organism="Glycine max"			
	/db_xref="taxon:3847"			
	/clone="GENOME SYSTEMS CLONE ID: Gm-cl012-871"			
	/clone_id="Gm-cl012"			
	/tissue_type="Apical shoot tips, 9-10 day old etiolated seedlings"			
	/lab_host="XL10-Gold"			
	/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from the apical shoots of 9 to 10 day old etiolated seedlings. The shoot tips including any emerged leaves were harvested for mRNA isolation. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."			
BASE COUNT	153 a 105 c 142 g 138 t			
ORIGIN				
Query Match	2.9%	Score 197;	DB 62;	Length 538;
Best Local Similarity	73.6%;	Pred. No. 2.4e-28;		
Matches 251;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;
6336	ttaggttaaaagccttttgaagaagaaagtgtgagatgcgtgcgactcgtactgca	6395		
196	TTGGGTAAAGAGACTTCTTAAGAGAGAAAAGCTTGAAATGTTGGTAGATCTGATTTACA	255		
6396	gaacaattacattgacacagaagttgagcagcttattcaagatgacattacgtaccca	6455		
256	AACCAACTATATAGAAAGTGAAGTAGAACAAGTTAATCCAGGTTGACACTCTGCACACA	315		
6456	gggttgcccaatgggagcgccctaagtgtccagaggtgagccgaatgcttgaaggtgatg	6515		
316	AGGTTCCCGATGACCGACCTTAAGATGTCAGAAAGGTGAGAGATCTTGAAGGTGATGG	375		

QY	6516	ccttgagaagaagtgtgagcagctgtgcaaaaagtgtgaagtcattccatccaagacgtagaatt	6575
Db	376	CTTGGCGAGAAAGATGGATGAGTGGCAAAAGGTGGAAGTCTCCGGCAGAGAGTGGAGCT	435
QY	6576	agctccacatgcgaactctgcaatgtgatccctagactgcacagataactgtgactgtttga	6635
Db	436	GGGCCCTCATCTTAATTCGGATTGGATGTGTGCATCAACCGAAATCTGCATGCAGTAGA	495
QY	6636	attatctgtgtccaaataaacaagcagcataataatgtatgtaa	6676
Db	496	GTTATCTGTGTCCAGAGTGACCTTGGCACAGTAGTAATAATTA	536
RESULT	6		
LOCUS	AI727837	597 bp	mRNA
DEFINITION	BNLH15209 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase [Daucus carota], mRNA sequence.	EST	11-JUN-1999
ACCESSION	AI727837		
VERSION	AI727837.1	GI:5046689	
KEYWORDS	EST		
SOURCE	upland cotton.		
ORGANISM	Gossypium hirsutum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.		
AUTHORS	1 (bases 1 to 597)		
TITLE	Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.		
JOURNAL	ESTs from developing cotton fiber		
COMMENT	Unpublished (1999)		
	On Jun 5, 1998 this sequence version replaced gi:3187180.		
	Contact: Ben Burr		
	Biology Department		
	Brookhaven National Laboratory		
	Upton, NY 11973, USA		
	Tel: 516-344-3396		
	Fax: 516-344-3407		
	Email: burr@bnlxi1.bnl.gov		
	Seq primer: T3 Primer.		
FEATURES	Location/Qualifiers		
source	1..597		
	/organism="Gossypium hirsutum"		
	/cultivar="Acala Maxxa"		
	/db_xref="taxon:3635"		
	/clone_lib="Six-day Cotton fiber"		
	/tissue_type="Immature fiber"		
	/dev_stage="Six days post anthesis"		
	/lab_host="X11-Blue"		
	/note="Vector: pBluescript II KS+"		
BASE COUNT	165 a 107 c 165 g 148 t	12 others	
ORIGIN			
Query Match	2.9%: Score 191.8; DB 51; Length 597;		
Best Local Similarity	69.8%; Pred. No. 2.5e-27;		
Matches 259; Conservative	0; Mismatches 112; Indels 0; Gaps 0		
QY	5828	gattggccaactgagagagattgcactagatctcttaggggcctatctaattgcat	5887
Db	11	GATGGCCCTCAAGAAAAAGAAATTAACCTAGGTGCAGCATGCGGCTGGATACCTTCAT	70
QY	5888	gaccattgtatcccaagattatccatgcgcgtgtgtaaaagctgcgaatatctttgac	5947
Db	71	GAACATTTGCATCCTTAAGATCATTCAGAGATGTGAAGGCTGCAAATGATTAATGGAT	130
QY	5948	gaagaattcagagctgtgtcaggtgatttctgggttagctagctcatgtgattacaagat	6007
Db	131	GAACATTTTGAAGAGATGTGTGCGGTGACTTTGGCCTTGGGAAGTTGTGTGATGTAGGCGG	190
QY	6008	accacattacacgactgctgtaaggggtaccatctgggacacatagctcccgagctactcgtg	6067

Db 191 ACAGATGTGACACGTCAGTTCGGGACATGGGACATACACCCGAGTACTTGCC 250
Qy 6068 actggaagatcatcagaagaagaccgagctcttggtatggaatgctcccaagagcc 6127
Db 251 ACTGGGAAGTATCAGAAAGAACGAGTGTCTTGATGGATTTAGCTCTAGAGCTT 310
Qy 6128 attactgacagagagcctttgatctgctgccttcgacagatgatgattggtg 6187
Db 311 GTGACAGGTCAACGTGCATGATTCTCAGCGTTGGAAATGAGATGATGTCTGCTT 370
Qy 6188 ttgagttggt 6198
Db 371 CTGACTATGT 381

RESULT 7
LOCUS A1729440 703 bp mRNA EST 11-JUN-1999
DEFINITION BNGH13383 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar
to (U93048) somatic embryogenesis receptor-like kinase [Daucus
carota], mRNA sequence.
ACCESSION A1729440
VERSION A1729440.1 GI:5048292
KEYWORDS EST.
SOURCE upland cotton.
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 703)
REFERENCE Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.
ESTs from developing cotton fiber
Unpublished (1999)
On Jun 5, 1998 this sequence version replaced gi:3188791.
COMMENT
JOURNAL Contact: Ben Burr
Biology Department
Brookhaven National Laboratory
Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
Email: burrdnluxl.bnl.gov
Seq primer: T3 Primer

FEATURES
Source location/Qualifiers
1..703
/organism="Gossypium hirsutum"
/cultivar="Acala Maxxa"
/db_xref="taxon:3635"
/clone_lib="Six-day Cotton fiber"
/issue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="XLI-Blue"
/note="Vector: pBluescript II KS+"
BASE COUNT 188 a 148 c 180 g 186 t 1 others
ORIGIN

Query Match 2.9%; Score 191.8; DB 51; Length 703;
Best Local Similarity 69.8%; Pred. No. 2.4e-27;
Matches 259; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 5828 gattggccaactagaggagagattgcactgacatcttgaaggagcctatcaatgcat 5887
Db 11 GATTGGCCCTCAAAAGAAATACCTAGTGCAGATGTGGCTGAGTACCTTCAT 70
Qy 5888 gaccatttgatcccaaatatcatcgcgcatgtaaaagctgcaatatattggac 5947
Db 71 GACATTTGCAATCTTAATCATTCATGAGATGGAAGCTGCAATGATATTGAT 130
Qy 5948 gaagaattcgaagcgtctgtagtattggtgtagtagcctcatggaattcaagat 6007
Db 131 GAAAGATTGGAAGCAGTGTGCGTACTTTGGCCTTGGAAGTGTGTGATGTGAGCGG 190
Qy 6008 acccatgttaagactgctgtaagggtaccatttggcacataagctcccgagttactctg 6067

Db 191 ACAGATGTGACACGTCAGTTCGGGACATGGGACATACACCCGAGTACTTGCC 250
Qy 6068 actggaagatcatcagaagaagaccgagctcttggtatggaatgctcccaagagcc 6127
Db 251 ACTGGGAAGTATCAGAAAGAACGAGTGTCTTGATGGATTTAGCTCTAGAGCTT 310
Qy 6128 attactgacagagagcctttgatctgctgccttcgacagatgatgattggtg 6187
Db 311 GTGACAGGTCAACGTGCATGATTCTCAGCGTTGGAAATGAGATGATGTCTGCTT 370
Qy 6188 ttgagttggt 6198
Db 371 CTGACTATGT 381

RESULT 8
LOCUS A1486453 471 bp mRNA EST 29-JUN-1999
DEFINITION EST244774 tomato ovary, TMU Lycopersicon esculentum cDNA clone
CLED8G2, mRNA sequence.
ACCESSION A1486453
VERSION A1486453.1 GI:4381824
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 471)
REFERENCE Alcalá, J., Vrebalov, J., White, R., Matera, A.L., Viston, T.,
Holt, I.E., Liang, F., Upton, J., Rønning, C.M., Craven, M.B.,
Fujii, C.Y., Bowman, C.L., Niernann, W., Fraser, C.M., Venter, J.C.,
Martin, G.B., Tanksley, S.D. and Giovannoni, J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
On May 7, 1998 this sequence version replaced gi:3121325.
COMMENT
JOURNAL Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.

FEATURES
Source location/Qualifiers
1..471
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_lib="CLED8G2"
/clone_lib="tomato ovary, TMU"
/issue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLED - Tomato Carpel EST Library. Oligo(dT)-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."

BASE COUNT 136 a 88 c 112 g 135 t
ORIGIN

Query Match 2.7%; Score 184; DB 47; Length 471;
Best Local Similarity 66.8%; Pred. No. 8.1e-26;
Matches 262; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 5801 gagctcagccatcagaactcccccctgattggccaactagaggagagattgactaaga 5860
Db 54 GAACCTAACTGGGGAGCCTGTTTGAATGGCCCACTAAGAACGCTGTGCGGTGGT 113
Qy 5861 tctctaggggcctatcaaatattgcatgacattgtagtcccaagattatcatcgcat 5920

Db 114 ACTGCACGTGACTGATACCTACATACACCTGTAATCCAAAGATTATTACCGCGAT 173
 QY 5921 gtaaaagctgcaaatatattattgacgaagaatttgagctgtgttagtgattcgg 5980
 Db 174 GTTAAAGCAGCTAATGTGTACTAGATGAAGTTTGAAGCTAGTGGATGGC 233
 QY 5981 ttagtaggctcagtgattacaagaataccatgttaccagctgtgaagggtaccatt 6040
 Db 234 CTGGCAACCTAGTGTACCTTGAAGAAACCAATGTGACTACTCAAGTTTCGTGATCAATG 293
 QY 6041 gggacacatagctcccgagactctcgcagtgaaagtcataagaagaagcagatgcttt 6100
 Db 294 GGCATATAGCTCTCTAATACCTATCCACTGCAATATCATGAAAAACATGATTTT 353
 QY 6101 ggtttaggataatgctcctcagaagctcattacagcagaaggctttgtatctgtcgc 6160
 Db 354 GGCTATGGAATCATGCTTTTGGAAATTGTAACCGGCCAACGTCAATGACTTCTCAGCG 413
 QY 6161 ctggcgaacgatgatgatgttatgtgttgg 6192
 Db 414 CTAGAAGATGAAGATGATGTCTTGTGCTCGA 445

RESULT 9
 A115643 380 bp mRNA EST 03-DEC-1998
 LOCUS A081P60u Hybrid aspen plasmid library Populus tremula x Populus
 DEFINITION tremuloides cDNA 5', mRNA sequence.
 ACCESSION A115643
 VERSION A115643.1 GI:3856928
 KEYWORDS EST.
 SOURCE Populus tremula x Populus tremuloides.
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicotyledons; Magnoliophyta; eudicotyledons; core
 eudicotyledons; Rosidae; eurosids I; Malpighiales; Salignaceae; Populus.
 1 (bases 1 to 380)
 Sterky,F., Regan,S., Karlsson,J., Hertzberg,M., Rohde,A.,
 Holmberg,A., Amlin,B., Bhalerao,R., Larsson,M., Villarroel,R., Van
 Montagu,M., Sandberg,G., Olsson,O., Teerl,T.T., Boerjan,W.,
 Gustafsson,P., Uhlen,M., Sundberg,B. and Lundberg,J.
 Gene discovery in the wood-forming tissues of Poplar: Analysis of
 5,692 expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)
 99007314
 On Jan 17, 1998 this sequence version replaced gi:2044473.
 CONTACT: Sterky F
 Department of Biotechnology
 Royal Institute of Technology (KTH)
 Teknikringen 34, S-100 44 STOCKHOLM, Sweden
 Tel: +46 8 790 8287
 Fax: +46 8 24 54 52
 Email: fredrik@biotech.kth.se
 PCR PRIMERS
 FORWARD: AAAGGGGATGCTGCTGCAAGCG
 BACKWARD: GCTTCGGCTGATGTGTGTG
 Seq primer: CCGTGAACGACGCGCAG
 High quality sequence stop: 380.
 Location/Qualifiers
 1..380
 /organism="Populus tremula x Populus tremuloides"
 /db_xref="taxon:47664"
 /clone_lib="Hybrid aspen plasmid library"
 /tissue_type="Cambial region"
 /dev_stage="1.5 m actively growing tree"
 /lab_host="E.coli"
 /note="Vector: pBluescript SK; Site.1: SalI; Site.2: NotI;
 Cambial region tissues, including developing xylem, the
 meristematic cambial zone and the developing and mature
 phloem, was harvested from 1.5 m actively growing trees.
 cDNA was prepared and cloned into lambda gt22a. DNA was
 isolated and subcloned into pBluescript SK using SalI and

BASE COUNT 109 a 67 c 99 g 102 t 3 others
 ORIGIN

Query Match 2.7%; Score 183.8; DB 43; Length 380;
 Best Local Similarity 80.1%; Pred. No. 8.9e-26;
 Matches 226; Conservative 0; Mismatches 55; Indels 1; Gaps 1;

QY 5918 gatgaaagctgcaaatatattattgacgaagaatttgagctgtgttagtgattt 5977
 Db 1 GATGTTAAAGCTGCAATATTTGTGTGACAGAGAAATTCAGGWTGATGCGGGATTTT 60
 QY 5978 gggtagtaggctcagtgattacaagaataccatgttaccagctgtgaagggtacc 6037
 Db 61 GGGCGNCAAAATTTATNACTACAGAGATACGATATCCACTGCTGT-CGTGTATCA 119
 QY 6038 attggacacatagctcccgagactctcgcagtgaaagtcataagaagaagcagatgct 6097
 Db 120 ATAGGACACATGCGTCCAGAGTACTACTACTGTAATCATCTGAAACATGATGTT 179
 QY 6098 ttggattagggataatgctcctcagaagctcattacagcagaaggctttgtatctgtc 6157
 Db 180 TTTGGGATGGAATCATGCTTCTGAGCTAATTAATCTGACAGAGGGCTTTCGACTTGCC 239
 QY 6158 cgccttgcaacgatgatgatgttatgtgttggattggta 6199
 Db 240 CCGCTGGCAATGATGACATGATGATTTACTTGAATGGGA 281

RESULT 10
 A2329855 475 bp mRNA EST 28-JAN-2000
 LOCUS A2329855 rootphos(-) Medicago truncatula cDNA clone MHRP-24H06,
 DEFINITION mRNA sequence.
 ACCESSION A2329855
 VERSION A2329855.1 GI:6800450
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
 Medicago.
 1 (bases 1 to 475)
 Harrison,M.J., Liu,J., Gonzales,M.B. and Ellis,L.
 ESTs from phosphate starved roots
 Unpublished (1999)
 On May 1, 1997 this sequence version replaced gi:2059389.
 CONTACT: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 Developmental stage: phosphate starved; Date: 01/23/00; Submitted
 to the Database of Expressed Sequence Tags (dbEST) on 01/26/00;
 More information is available at
 'http://chryslr.tamu.edu/medicago'.
 Seq primer: 73.
 Location/Qualifiers
 1..475
 /organism="Medicago truncatula"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone_lib="MHRP-24H06"
 /tissue_type="rootphos(-)"
 /clone_lib="rootphos(-)"
 /note="Vector: pBluescript SK -; Site.1: EcoRI; Site.2:
 XhoI; cDNA was prepared from polyA+ enriched RNA. The
 cDNA was directionally ligated into the Unizap XR vector
 from Stratagene and packaged using GigaPack III Gold

FEATURES
 source

FEATURES
 source

Ibaraki,
 Japan 305
 Tel: 0298-38-7441
 Fax: 0298-38-7468
 Email: minob@rics01.riken.go.jp
 PROJECT="RCP"
 Location/Qualifiers
 1. .299
 /organism="Oryza sativa"
 /strain="Nipponbare, sub-species Japonica"
 /db_xref="taxon:4530"
 /clone="R2976.1A"
 /clone_lib="Rice root"
 /note="Prepared from seedling root."

BASE COUNT 93 a 54 c 67 g 85 t
 ORIGIN

Query Match 2.6%; Score 175.4; DB 20; Length 299;
 Best Local Similarity 74.4%; Pred. No. 3.8e-24;
 Matches 221; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Oy 5846 aggaattgactagagctctctcgaaggccatctaatgcatgacccatgtgacccaag 5905
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 3 AGGATCGCAGCAGCTTCGCCAGGGGGCTGCTATTACATGATCATGCCAACCAAG 62
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 5906 attaccatcgcatgttaagaagctgcaaatattattgtagagcaagattggagctgtt 5965
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 63 ATCATCCATTCGTGATGTCAGAGCGGCAAAATTTATTAATGATGAAGCTTGAGCTGA 122
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 5966 gtaaggtagatttggttagctagagctcattgattacaagaataccatgttagactgtc 6025
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 123 TTGGAGATTGTTGGTGGCCAAACCAATGGATTACAAAGATACCATGTAACAACTGCA 182
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 6026 gtaagggttacattggcgcacatagctcccgagactctcgactggaagatcatagag 6085
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 183 GTCATGCAATGATGGCATATTCGACCAAGATATCTTTCACAGGAAATTCCTCTGAG 242
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 6086 aagaccagatgtcttggtagtgggaataatgctcctagagctcattactgagcagag 6142
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 243 AAAACTGATGTATTTGGCTATGATGATCATGCTTTGGAGCTTATTACAGGAAAAACG 299
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
 B27171 464 bp DNA GSS 13-OCT-1997
 LOCUS B27171 TAMU Arabidopsis thaliana genomic clone T2E10, genomic
 DEFINITION survey sequence.
 ACCESSION B27171
 VERSION B27171.1 GI:2513137
 KEYWORDS GSS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsis.
 1 (bases 1 to 464)
 Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and
 Venter,J.C.
 TITLE Use of a BAC End Sequence Database To Identify Minimal Overlaps for
 Arabidopsis Genomic Sequencing
 JOURNAL Unpublished (1997)
 COMMENT Other_GSSs: T2E10TR
 Contact: Steve Rounsley
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: rounsley@igf.org
 Seq primer: M13-21
 Class: BAC ends

High quality sequence stop: 464.
 Location/Qualifiers
 1. .464
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="T2E10"
 /clone_lib="TAMU"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; Site_1: HindIII; Site_2:
 HindIII; Produced by Rod Wing"

BASE COUNT 119 a 87 c 115 g 143 t
 ORIGIN

Query Match 2.6%; Score 173; DB 84; Length 464;
 Best Local Similarity 65.6%; Pred. No. 1.1e-23;
 Matches 292; Conservative 0; Mismatches 135; Indels 18; Gaps 2;

Oy 4728 ttcccttct 4787
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 20 TTAACATATTTGGTAAATAAATTAATTAAGGTGGTATGATGCTACTGAGCCATTGGCG 79
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 4788 gggagtagctgctgctgct 4847
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 80 GAGAGTTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 139
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 4848 ggggagagaagaacgcgaggaacattctcttgatgctccaggttagctctgtaataga 4907
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 140 GCGGTAGAGAAACCTCAAGATTCCTTTGATGATTCCTGATGATGATGATGATGATGATGAT 199
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 4908 tatctattgaagcgctac-----tgctgtagacttgcttctctctctctctctctct 4955
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 200 AATATCAAGCTTGTTCATTTCAATTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTC 259
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 4956 agttaactcagctggaagagccagagatgcacttgctgaactgaagaggtttctc 5015
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 260 TCTGTGATGACGCGAGAGAGACCTGAGGCTTCATCTGGGAGCTTAAGCGTTCCTCTC 319
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 5016 tgcgaattgcaagctgcagacgatactttag-----taccatcttggaagagtg 5069
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 320 TACGGAACTCAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 379
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 5070 gatttgtaaggtgtataagggagcgcctgctgtagtgctcacttgtagcagtttaagagc 5129
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 380 GGTTCGAAAGTCTACAAATGCGCTTCCTGATGGAACACTGTTCGACGCAAAACGCG 439
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 5130 ttaagaagagacgacacacccaggtg 5154
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 440 TTAAAGAAAGACGAAACCCCAAGTGG 464
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
 AA738544 991 bp mRNA EST 17-AUG-1998
 LOCUS AA738544 SBLKX Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
 DEFINITION SBLKX similar to protein kinase, highest similarity to
 receptor-like protein kinases, mRNA sequence.
 ACCESSION AA738544
 VERSION AA738544.1 GI:3421469
 KEYWORDS EST
 SOURCE sorghum.
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eumhyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
 Poaceae; Sorghum.
 1 (bases 1 to 991)
 Annen,F., Chang,J.-L., Paterson,A.H. and Stockhaus,J.
 TITLE Characterization of 14 different putative protein kinase cDNA
 clones of the C4 plant Sorghum bicolor
 JOURNAL Mol. Gen. Genet. 259 (1), 115-122 (1998)
 MEDLINE 98409267
 COMMENT On Nov 29, 1993 this sequence version replaced gi:636080.

Contact: Stockhaus J
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49)-211-81-14953
Fax: (49)-211-81-14871
Email: stockhaeuni-duesseldorf.de

Insert Length: 1500 Std Error: 0.00.

FEATURES

source
1. .991
/organism="Sorghum bicolor"
/strain="cv. TX430"
/db_xref="taxon:4558"
/clone="SBRK2"
/clone_lib="Sorghum bicolor cv. TX430 leaf"
/dev_stage="green"
/note="Organ: leaf; Vector: Lambda ZAP II; Site_1: Eco RI;
Site_2: Xho I; leaf cDNA library from green leaves,
unidirectionally cloned"
BASE COUNT 253 a 194 c 293 g 251 t
ORIGIN

Query Match 2.6%; Score 172.2; DB 38; Length 991;

Best Local Similarity 68.0%; Pred. No. 1.5e-23;

Matches 240; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

OY 5828 gattggccaactagagagagattgcacatgacatctctcagggccatcaattgcat 5887
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 358 GATTGGCTGCAAGAGACGAGTGGCTATAGGCACAGCTCGTGAGACTGAGATTTGGAC 417
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5888 gaccattgtatcccaagattatccatcgcatgtataaagctgcgaatatattatggac 5947
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 418 GAGCACTGCATCTTAAGATTATACATCGACGTCAAGCTGCCAATGCTTGCTGAT 477
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5948 gaagaattggagctgttgaattgattgggttagctcagagctcattgattacaaggat 6007
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 478 GAAGTTTGAACCGGTTGTTGGATTGGCGTTGGCCAACTGGTGTATACAGAG 537
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 6008 acccatgttaccgactgtctgaagggttacattgggacacatagctccgagatcctcg 6067
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 538 ACATCTGTGACTACTCAGTCCGCGGAAACATGGTCAATGCCCCCTGAAATTTGTC 597
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 6068 actgaaagtcacagaaagacccgattgtctgttattgagataatgtctcctagagctc 6127
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 598 ACTGGGAAGTATCCGAGAGAACCGATGTTTGGTTACGGCATATGCTTTCCAGCTA 657
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 6128 attactgacagagagctttgattctgtcgccttcgcgaacgatgatgt 6180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 658 GTCACTGGTCAGCGCCATAGACTTTTCAGTCTGGAGAGAGAGATGATGT 710
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15

AA738545

LOCUS

DEFINITION

AA738545 1435 bp mRNA EST 17-AUG-1998
SBRK3 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
SBRK3 similar to protein kinase, highest similarity to
receptor-like protein kinases, mRNA sequence.

ACCESSION

AA738545

VERSION

AA738545.1

KEYWORDS

EST.

ORGANISM

Sorghum

Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;

Poaceae; Sorghum.

REFERENCE

1 (bases 1 to 1435)

AUTHORS

Annen, F., Chang, J.-L., Paterson, A. H. and Stockhaus, J.

TITLE

Characterization of 14 different putative protein kinase cDNA

clones of the C4 plant Sorghum bicolor

Mol. Gen. Genet. 259 (1), 115-122 (1998)

MEDLINE 98409267
On Nov 29, 1993 this sequence version replaced gi:636085.

COMMENT

Contact: Stockhaus J
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49)-211-81-14953
Fax: (49)-211-81-14871
Email: stockhaeuni-duesseldorf.de

Insert Length: 1900 Std Error: 0.00.

FEATURES

source
1. .1435
/organism="Sorghum bicolor"
/strain="cv. TX430"
/db_xref="taxon:4558"
/clone="SBRK3"
/clone_lib="Sorghum bicolor cv. TX430 leaf"
/dev_stage="green"
/note="Organ: leaf; Vector: Lambda ZAP II; Site_1: Eco RI;
Site_2: Xho I; leaf cDNA library from green leaves,
unidirectionally cloned"
BASE COUNT 365 a 300 c 389 g 381 t
ORIGIN

Query Match 2.5%; Score 170.6; DB 38; Length 1435;

Best Local Similarity 67.7%; Pred. No. 2.9e-23;

Matches 239; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

OY 5828 gattggccaactagagagagattgcacatgacatctctcagggccatcaattgcat 5887
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 802 GATTGGCTGCAAGAGACGAGTGGCTATAGGCACAGCTCGTGAGACTGAGATTTGGAC 861
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5888 gaccattgtatcccaagattatccatcgcatgtataaagctgcgaatatattatggac 5947
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 862 GAGCACTGCATCTTAAGATTATACATCGACGTCAAGCTGCCAATGCTTGCTGAT 921
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 5948 gaagaattggagctgttgaattgattgggttagctcagagctcattgattacaaggat 6007
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 922 GAAGTTTGAACCGGTTGTTGGATTGGCGTTGGCCAACTGGTGTATACAGAG 981
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 6008 acccatgttaccgactgtctgaagggttacattgggacacatagctccgagatcctcg 6067
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 982 ACATCTGTGACTACTCAGTCCGCGGAAACATGGTCAATGCCCCCTGAAATTTGTC 1041
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 6068 actgaaagtcacagaaagacccgattgtctgttattgagataatgtctcctagagctc 6127
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1042 ACTGGGAAGTATCCGAGAGAACCGATGTTTGGTTACGGCATATGCTTTCCAGCTA 1101
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 6128 attactgacagagagctttgattctgtcgccttcgcgaacgatgatgt 6180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1102 GTCACTGGTCAGCGCCATAGACTTTTCAGTCTGGAGAGAGATGATGT 1154
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: June 23, 2000, 19:05:52
Job time: 27385 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 23:30:33 ; Search time 12463.9 Seconds

(without alignments)
-141.659 Million cell updates/sec

Title: US-09-180-798-2

Sequence: 1815
1 gacaaataccatggaatat.....ttaaaaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY_MNC
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p12:*
10: gb_p12:*
11: gb_p12:*
12: gb_p12:*
13: gb_p12:*
14: gb_p12:*
15: gb_p12:*
16: gb_p12:*
17: gb_p12:*
18: gb_p12:*
19: gb_p12:*
20: gb_p12:*
21: gb_p12:*
22: gb_p12:*
23: gb_p12:*
24: gb_p12:*
25: gb_p12:*
26: gb_p12:*
27: gb_p12:*
28: gb_p12:*
29: gb_p12:*
30: gb_p12:*
31: gb_p12:*
32: gb_p12:*
33: gb_p12:*
34: gb_p12:*
35: gb_p12:*
36: gb_p12:*
37: gb_p12:*
38: gb_p12:*
39: gb_p12:*
40: gb_p12:*
41: gb_p12:*
42: gb_p12:*
43: gb_p12:*
44: gb_p12:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_htg4:*
50: em_htg5:*
51: gb_p13:*
52: gb_p13:*
53: gb_p13:*
54: gb_p13:*
55: gb_p13:*
56: gb_p13:*
57: gb_p13:*
58: gb_p13:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1815	100.0	5 A67797	A67797 Sequence 2
2	1755	96.7	8 DCU93048	U93048 Daucus caro
3	934.8	51.5	5 A67827	A67827 Sequence 32
4	560.2	30.5	5 A67796	A67796 Sequence 1
5	434	23.9	8 F23M19	AC007454 Arabidops
6	398	21.9	7 ATEF17M5	AL035678 Arabidops
7	395.6	21.8	8 AC012654	AC012654 Arabidops
8	372.4	20.5	50 AC006436	AC006436 Arabidops
9	317.6	17.5	8 SBRXL1	Y14600 Sorghum bic
10	270.8	14.9	5 A67815	A67815 Sequence 20
11	252.6	13.9	7 AB013395	AB013395 Arabidops
12	246.8	13.6	7 AB025639	AB025639 Arabidops
13	216.2	11.9	8 F17123	AF160182 Arabidops
14	215.4	11.9	8 AC002292	AC002292 Genomic s
15	209.4	11.5	50 AC005170	AC005170 Arabidops
16	203	11.2	7 AB005234	AB005234 Arabidops
17	200.8	11.1	7 AB012245	AB012245 Arabidops
18	178.4	9.8	7 AB007644	AB007644 Arabidops
19	164.6	9.1	8 AC002330	AC002330 Arabidops
20	163.8	9.0	50 AC005312	AC005312 Arabidops
21	161.4	8.9	7 RICEPK1	L27821 Oryza sativ
22	160.8	8.9	8 ATEF884	AL034567 Arabidops
23	160.8	8.9	8 ATEF10M6	AF024648 Arabidops
24	158.4	8.7	8 ZMA01016	AF024648 Arabidops
25	157.8	8.7	8 ZMA01016	AF024648 Arabidops
26	153.8	8.5	8 ZMA01016	AF024648 Arabidops
27	153.6	8.5	8 ZMA01016	AF024648 Arabidops
28	153.2	8.4	8 ATEF884	AF078082 Phaseolus
29	153.2	8.4	8 ATEF884	AF078082 Phaseolus
30	153.2	8.4	8 ATEF884	AF078082 Phaseolus
31	149	8.2	50 AC008262	AC008262 Genomic s
32	149	8.2	8 ATEF884	AF078082 Phaseolus
33	145.6	8.0	8 ATEF884	AF078082 Phaseolus
34	145.6	8.0	8 ATEF884	AF078082 Phaseolus
35	138.4	7.6	5 E12705	U47029 Arabidops
36	138.4	7.6	8 ATEF884	AF078082 Phaseolus
37	138.4	7.6	8 ATEF884	AF078082 Phaseolus
38	133.8	7.4	8 ATEF884	AF078082 Phaseolus
39	133	7.3	7 AB016890	AB016890 Arabidops
40	131.4	7.2	50 AC004684	AC004684 Arabidops
41	130.6	7.2	8 AC002392	AC002392 Arabidops
42	129.8	7.2	7 ATEF884	AF078082 Phaseolus
43	129.8	7.2	7 ATEF884	AF078082 Phaseolus
44	128.2	7.1	7 AB025636	AB025636 Arabidops
45	128.2	7.1	8 AF058826	AF058826 Arabidops

ALIGNMENTS

RESULT 1
LOCUS A67797 1815 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 2 from Patent WO9743427.
ACCESSION A67797
VERSION A67797.1 GI:4756623
KEYWORDS
SOURCE
ORGANISM
Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asterales; euasterids II; Apiales; Apiales; Daucus.
REFERENCE
1 (bases 1 to 1815)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICTIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
SOURCE location/Qualifiers
1..1815
/organism="Daucus carota"
/db_xref="taxon:4039"
94..1755
/note="unnamed protein product"
/codon_start=1
/protein_id="CABA2241.1"
/db_xref="GI:4756624"
/translation="MNRNSINILNMYQFDAYLDKYGVLTLELYSNNISGPIPSDLG
NLNLVSLDLYMNSFSGIPDITLKLRLRFLRLNNLSLGPIMSLTNTITLQVLDL
SNRLSPVPDNGSFLFTPISEFANLNLCGPVGRCPSPSPSPSPSPSPSPSPSP
PGONGPTGALGVAAAGALLFAPAPAFAMMRBRKREHFVPAEEDVEVLGOLK
RFSRLQVATDPTSTILRGGEKGYKGLADGSLVAAYVRLKEBRPGSELFOREV
EMIMAYHNRILRRGRCMPTPELILVYPMANGSVASCIKEOPSPSPSPSPSPSP
ALGSRKSTIHDRCDKTIIHRDYKANILIDEFEAVYVDEGLARLMDKDHHTTA
VRGTGLTAPETLSTGKSSSEKTVFEGIMLELTGQRAFDLARLNDVDMLDLYA
KSLRKERLEMLVDPDLENNYIDTEVBOILQVALLCTQSGSPMRPRKSEVYVRLLEDG
LAEKWEDEKQEVTHQVDELAPHRSTSEMILSDNHLAEFLSGPR"
BASE COUNT 531 a 354 c 415 g 515 t
ORIGIN

Query Match 100.0%; Score 1815; DB 5; Length 1815;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1815; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gccaataccattgaatatattgaacctaataattagtagtgcagtttaattcaaa 60
DB 1 GACAAATACCATTTGAATTTGAACCTAATTAATTAGTAGTGAAGTTAAATTCAAA 60
QY 61 ccatttaatttactttaaaaaaataattctataatgcgaacgtaacagataataatla 120
DB 61 CTCATTATTTTACTTTAAAAAATAATTCTATATGAATCGAACGATATAATATATA 120
QY 121 aattacatgcagttcaatgcagttcaatgcagtaataatggggtctctatgcattggag 180
DB 121 AATTATATGCAATTCACGTATCTTACCTTGACAAATATGGGGTCTTATGACATTTGAG 180
QY 181 cttaacagaataaataagtaggaacattccttagtagtctctgggaatctgacaatttg 240
DB 181 CTTTACAGCAATTAAGTAGTGACCAATTCCTATGATCTTGGGAATCTGACAAATTTG 240
QY 241 gtgagcttgaccctatacagaatagctctctggaacataccggacacattaggaag 300
DB 241 GTGAGCTTGACCTATACATGATAGCTCTCTGAGACCATACCGACACATTAGGAAG 300
QY 301 cttaacagaatgaattctctgctctcaacaaacaaacagctctctggtccaaattccaatg 360
DB 301 CTTACAGGCTTAAGATTCTTGCTCTCAACAACAACAGCCTCTGTTGTTCAATTCCAATG 360
QY 361 tcaactacataatacactctcaagtcctggatttacaacacacggtatcaga 420
DB 361 TCACCTACATATATATCAACTCTTCAAGTCCGTGATTTATCAAAACATCGGCTATCAGA 420

QY 421 ccagttaccggaataatggtcattcttcttcttcaacattacagtttgcacaaatttg 480
DB 421 CCAGTACCGGAATAAGGCTCATTTCTTTGTTTACACTATCACTTTGCCAATAATTG 480
QY 481 aattatctgagcccgtaacccgagagcccgccctgacatcccccatttgcacaca 540
DB 481 AATTATATGAGACCCGTAACCTGGAGGGCCCTGCGCTGATGTCCCATTTTGGCACCA 540
QY 541 cctccgttcatccacacatcaacagtagtaccgccccaagacaataatgltcccaatgagct 600
DB 541 CCTCCGTTTCATCCACCATCAACAGTACAGCTCCAGACAAAATGTTCCACATGGAGCT 600
QY 601 attgcttgagagatgagctgctgctgctgcttactggttctgtagcctggaatgacatt 660
DB 601 ATTGCTGGGGAGTAGCTGCTGCTGCTGCTGCTTACTGTTGCTGCACTGCAATGGCAATT 660
QY 661 gcatgtgctgagagagaacacccgagagaacattcttcttgatgctgagcagctgaagagac 720
DB 661 GCATGTGCTGAG 720
QY 721 ccagaatgacattctgtagcagttcaatgaagaagttctctctgagagaattgcaagtcgaag 780
DB 721 CCAGAATGTCACCTTGCTGTCACACTGAAGAGGTTTCTCTGCGAGAAATGCAAGTGGCAAG 780
QY 781 gattacttttagtaccattacttggagaaggtgatttggtaaggtgataagggagcgctt 840
DB 781 GATTACTTTTAGTACCATACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 gctgatgctcactctgtagcagtttaaaagcttaaaagaacgaacacgaagtgtgtgag 900
DB 841 GCTGATGCTCCTCTGTGACAGTATTAAGAGGCTTAAAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 ctggagattcaacagagagtggaatgtagttagcagtgctgagcagtcgaacatctctcgt 960
DB 901 CTGCAATTTCAACAGAGGTGGAATGATTCAGAGGCTGTCATCAATCTTCTGCGT 960
QY 961 ctacgtgcttctgtagcagcacaacacgaagcgcctctctgatatccatatactgataat 1020
DB 961 CTACGTGCTTCTGATGATGACACCAACAGAGGCGCTTCTTATATCCATACATGGCTAT 1020
QY 1021 ggaagtgttgcgtcgtgtttaaagagagcgtcaagccataagaactccctgattggcca 1080
DB 1021 GGAAGTGTGCTGCTGCTGTTTAAAGAGAGGCTCAGCATAGAACCTCCCTGATTGGCCA 1080
QY 1081 actagaagaagattgacatgaatcctgaagcctgaagggccttctatctgtagcagcattgt 1140
DB 1081 ACTAGAGAGAGATTGCTACTAGGATCTGCTAGGGGCTTCTTATTTGACATGACCATGT 1140
QY 1141 gatcccaagattatccatcgtatgataaagctgtaaaatataattatggacgaagaattt 1200
DB 1141 GATCCCAAGATTATCCATGATGATTAAGCTGCAAAATATATATTGAGAGAGAAATTT 1200
QY 1201 gaggctgttgtaggtatttggttagctagcgtcagatgattacaagaatatacattgtt 1260
DB 1201 GAGGCTGTGTAGGTATTTGGTTAGCTAGGCTCATAGGATTAAGAGATACCAATTT 1260
QY 1261 acaactgctgaagaggtacatttgggtacatagctcccgagttacccctgactgaggaag 1320
DB 1261 ACAACTGCTGTAAGGGGTACTTGGGCTACATAGTCCCGAATACTCTCGACCTGGAAG 1320
QY 1321 tcatcagaagaacagatgcttcttggttataggtatagctctttagagctcaactactgga 1380
DB 1321 TCATCAGAGAGAGCCGATGCTTGTGTTATGGGATTAAGCTCTTGGAGCTCATTAACCTGGA 1380
QY 1381 cagaagcctttgattctgctgctccttggagacagatagatgattatgttcttggattgg 1440
DB 1381 CAGAGAGCTTTTATCTTGTCTGCTGCTGCGAAGCATGATGATGATGATGATGATGAT 1440
QY 1441 gttaaaagccttttgaagaagaagaagttgagatgctgctgacatcctgacactggaagac 1500
DB 1441 GTTAAAAGCCTTTTGAAGAAGAAAAGTTGGAGATGCTGCTGATCTGACCTGAGAGAAC 1500
QY 1501 aattacattgacacagaagttgagcagcttattcaagtagcatctctgtacccaaggt 1560

|||||
Db 1501 AATTACATTGACAGAAAGTTGAGCAGACTTATTCAGTAGCATCTGTGATCCAGGGT 1560
QY 1501 tcgccaatgagcgccctaagatgctcagagtgatccgaatgcttaagtgatgctt 1620
Db 1501 TCGCAATGAGCGCGCTTAAGATGTCAGAGTAGTCGAAATGCTTAAGATATGCGCTT 1620
QY 1621 gcaagaagatgagcgagtgcaaaaagtagaagtcacatccatcaagcagtagaatgct 1680
Db 1621 GCAGAAAAAGTGGAGCGAGTGGCAAAAAGTAGAATCATCATCAACAGCTAGATTAGT 1680
QY 1681 ccaatcgaaattctgaatgagctcctagactcgacagataactgcatgctttgaatta 1740
Db 1681 CCACATCGAATCTTGATGATGATCCTAGACTCGACAGATTAACCTTGATGCTTTGAATTA 1740
QY 1741 tctgtccaaagataaagcagcatataaagtgaatgaattatctttttttgttaaa 1800
Db 1741 TCTGTCCAAAGATTAACAGCATATAAATGATGATTAATTTTATTTGTTAA 1800
QY 1801 aaaaaaaaaaaaaa 1815
Db 1801 AAAAAAAAAAAAAA 1815

RESULT 2
LOCUS DCU93048 1755 bp mRNA PLN 28-JUN-1997
DEFINITION Daucus carota somatic embryogenesis receptor-like kinase mRNA,
complete cds.
ACCESSION U93048
VERSION U93048.1 GI:2224910
KEYWORDS
SOURCE .
ORGANISM
Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids II; Apiales; Apiales; Daucus.
REFERENCE
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
TITLE A leucine-rich repeat containing receptor-like kinase marks somatic
plant cells competent to form embryos
JOURNAL Development 124 (10), 2049-2062 (1997)
MEDLINE 97313247
REFERENCE
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.
TITLE Direct Submision
JOURNAL Submitted (12-MAR-1997) Molecular Biology, Agricultural University
of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
FEATURES
Source
1..1755
/organism="Daucus carota"
/db_xref="taxon:4039"
34..1695
/note="SERK"
/codon_start=-1
/product="somatic embryogenesis receptor-like kinase"
/protein_id="AA661708.1"
/db_xref="GI:2224911"
/translation="NNRNSINILNQFIDAVLDKYVLTLELXSNISGPIPSDLG
NLTVLSIDLYNNSFSGPIPDITGLRLRLNNNSLSPISPLNITTLAYLD
SNRRLSGPVNDGSPSLFPISEFNNILNCGPVGPCGSPFPPPTPLPVOP
PGONGPTGAIAGVAGAAALFEAPAAAMWRNRPRRPFDAEDPEVHSGOLK
RSLRLQVATDPTFSTLIGGGRKYVRIADGSIVANKRKERTGGELQPTQV
EMISMAVHRLILRGPCMPTRRLVYYTMANGSASLIRRGQSEPLDPTKRI
ALSGARSLSLHDHDPKILIRHYKANIILDBEFAVVDGLRLNDYKDTHTTA
VRGLGTLAEPIYSLTSGSKSEKIDVFYGLMLLELTGGAFDLARLADDVMLDW
KSLRKERKLEMLVDPLENNYIDTEVEOLIOVALICTGSPERKMSBYVMLEGD
LAERKMDMORVEVYIHQDVELAPRHSWILDSIDNLHAFELSGPR"
BASE COUNT 506 a 347 c 407 g 495 t
ORIGIN
Query Match 96.7%; Score 1755; DB 8; Length 1755;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1755; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 cttatttaatttactcttaaaaaataattctatatgaaatcgtlaacagtaataatttta 120
Db 1 CTCATTAAATTTTACTTTAAAAAATAATTCCTAATGAATCTTAACAGATTAATATTA 60
QY 121 aattacatgagctcagctatgcttacccttgacaataatggggtctcatgacttgag 180
Db 61 AATTACATGAGTTCACCTATGCTTACCTTGACAAATATGGGGTCTTATACATTGGAG 120
QY 181 cttaacagcaataaacaataagtgaccaattccctagtgactcttggaactcgtacaattg 240
Db 121 CTTTACAGCATTAATTAAGTGGACCAATTCCTAGTGAATCTGGGAATCTGACAAATTTG 180
QY 241 gtagactgagctatfacatgaaatagctctctcttgaaactatccggaaacttgagaag 300
Db 181 GTGAGCTTGACCTATACATGAAATAGCTTCTTGAACTATACCGACACATTGGAAG 240
QY 301 cttaacaggttaagattctgctctcaacaacaacagcctctctgccaattccaatg 360
Db 241 CTTACAAAGGCTTAAGATTTGCGTCTCAACAACAACAGCCTCTGTCGAATTCGAATG 300
QY 361 tcaatgactaatatcaactctcaagctcgtgattatataaacaatcggtctatcgga 420
Db 301 TCACTGACTAATTTTCAACTCTTCAAGTCTGATTTATCAACAAATCGGCTATCGAGA 360
QY 421 ccagtaacggataatggtcattcttcttggtttacacatcattggttgcataatttg 480
Db 361 CCAATACGGATTAATGGCTCATTTCTTTGTTTACACCAATTCAGTTTGCCAAATTAATG 420
QY 481 aattatgtagccgttaacttgaggagccctgctgatatcccccaatttgccaaca 540
Db 421 AATTATATGTGACCGGTAATGGAAGCCCTGCCGATCCCAATTTCCGACACA 480
QY 541 cctccgttcatcccaaccatcaacaagtagacacctccaaggaacaatggtccactgagct 600
Db 481 CCTCCGTTATCCCAACATCAACAGTAGACCTCCAGGACAAATATGTCACATGAGACT 540
QY 601 atgagctggagaaatagctgctgctgctgttcttaccgtgttgcctgcgaactgagcttt 660
Db 541 ATTGCTGGGGAATAGCTGCTGTGCTGCTTACGTTGCTGCAACCTGCATATGGCATTT 600
QY 661 gcatggtgcgaggaagaacaacccgcagaaacattcttcttgatgctccagctggaagagac 720
Db 601 GCATGATGCGGAGAGAAAGAAACCGGAGAAACATTCTTTGATGATGCCAGTGAAGAGAC 660
QY 721 ccagaaatgcaacttgctcaactgaagaggttctctcgagaaatgcaagtgcgaacg 780
Db 661 CCAGAACTGCACCTTGCTCAACGAAAGAGGTTTCTGCGGAATTCGAAGTCCCAACG 720
QY 781 gatacttctgtaaccatcttggaagaggtgatttgtaaggtgtataaggaagacctt 840
Db 721 GATATCTTTAGTACATACTGTGGAAGAAGGTGATTTGTAAGTATTAAGAGACGCTT 780
QY 841 gctgagtgtctcaacttgtagcaggttaaaaggcttaagaaagaaacgaacacaggtgtgag 900
Db 781 GCTATATGCTCACCTTGAGCAGATTAAAGGCTTAAGAAGAAAGCAACACGAGTGTGAG 840
QY 901 ctgagagttcaacagaggtggaatgattagcagtgctgctcatcgaaactcttgctg 960
Db 841 CTGAGAGTTTCAACAGAGGTGGAATGATTACATAGCGTGTGCATCGAAATCTTGTGCT 900
QY 961 ctacgvtgcttctgcatgacaaccaacagagcggtctctgtatatacatcaatgctaat 1020
Db 901 CTACGTGTTTTCGATGACACCAACAGACGCGCTCTTGTATTCACATATGCTATAT 960
QY 1021 ggaagtcttgcgtggtttaaagagcgtcaagcaatgaagaccctccctgattgagca 1080
Db 961 GGAAGTCTTGCTGCTGTTTAAAGAGCGTCAACCATCAAGACCTCCCTTGATTGGCCA 1020
QY 1081 actagaagaagagatgactagaagatcgttaaggggcttcttatttgatgacattgtt 1140

Db 1021 ACTAGAAAGAGATGCTAGATCTGTAAGGCGCTTCTTATTGTCATGACATGT 1080
Qy 1141 gatcccaagattatccatctgtatgtaaagcgcgaatatatatattgacgaagattt 1200
Db 1081 GATCCCAAGATTATCCCTCGATGTAAAGCTGCATAATATATTATGGACGAATATT 1140
Qy 1201 gaagctgttaggtgtattctgtgttagtagcgcatacgtatatacagaagatacccatgtt 1260
Db 1141 GAGCGCTGTGTAGGTGATTTTGGGTATAGTACGCTCATGATTAACAAGATACCCATGTT 1200
Qy 1261 acaactctgttaaggggttaacctgtgggtacatagctcccgagtaacctctccgactgtaag 1320
Db 1201 ACAACTCTGTAAAGGGGTACTGTGGCTACATAGCTCCGAGTACCTCTCGACTTGAAG 1260
Qy 1321 tcacagaagaagacgcgtctctgtgtataggtatagctcttagagctcattactgtga 1380
Db 1261 TCATCAGAAGAAGACCGATGCTTTGGTATAGGATTAATGCTCTTAAGCCTACTACTGA 1320
Qy 1381 cagaagactttgattcttctgcctctgcgaacgatagtatgtatgtttagttgattg 1440
Db 1321 CAGAGACTTTTGTATCTTGCTCGCTTCCGACAGATGATGATGTATGTTGTGATGG 1380
Qy 1441 gttaaagccctttgaaagagaaagttgagatgctgtgactcctgacctgagaaac 1500
Db 1381 GTTAAAGCCTTTTGAAGAAGAAATTTGAGATGCTGTGCTCATCTGACCTGGAGAAC 1440
Qy 1501 aattacattgacacagaagttgagcagcttattcgaagtagcattactctgttaccaggg 1560
Db 1441 AATTACATTGACACAGAAGTTGAGCAGCTTATTCATACATACATTAATCTTACCCAGGGT 1500
Qy 1561 tcgcgaatgtggggggtcttaagatgtcagaggttagtcgaagctgttaaggttagtcctt 1620
Db 1501 TCGCCATGTGGCGGCTTAAGATGTCAAGGTACTCCGATGCTTGAAGTGTATGGCTTT 1560
Qy 1621 gcagaagaagtgagcagagtgcaaaagtagaagtagcatccatcaagaagcgtagaatagct 1680
Db 1561 GCAGAAAAGTGGAGAGTGGCAAAAAGTAGAAGTCAATCATCAAGCAGTAGAATAGCT 1620
Qy 1681 ccaatgtgaacttctgtaagtagcttagactgcgacataacttctgtttagtaata 1740
Db 1621 CCACATCAACTCTGTGATGTGATCTAGACTCGACAGATTAATCTGCTTTTGAATTA 1680
Qy 1741 tctgttccaagataaagcatalataaaatgtgaatgaataaattttttttagttaa 1800
Db 1681 TCTGTCCAAATTAACAGCATATAAATGATGAATTAATTTTATGATTAA 1740
Qy 1801 aaaaaaaaaaaaaa 1815
Db 1741 AAAAAAAAAAAAAA 1755

RESULT 3
LOCUS A67827 2089 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 32 from Patent WO9743427.
ACCESSION A67827
VERSION A67827.1 GI:4756649
KEYWORDS
SOURCE
ORANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsi.
1 (bases 1 to 2089)
De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
PRODUCTION OF APOMICRITIC SEED
Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
Location/Qualifiers
1..2089
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

CDS
195..2072
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42254.1"
/db_xref="GI:4756650"
/translation="MESSYVFILSLILPNHSLIMASANEEDALHTLAVTLVDPN
NVLQSDPPLVPCWTFHVTCCNNNSVIRVDGNALSGHLVPELGLVNLQYLEYS
LNITGPISNLGNLNLVSLDYLNSFSGPIPELSGLSLKSLFLNLNNSLGLSIPMS
LNITLLOVLDLSNNRLSGVDPNGSFLETPISFANNLDLCGPVTSHPGSPFSP
PPPIOPPPVSPISGYSIGTGAAGVGAALFPAAPAIAMFMRMRSPDIFDVA
EBDPYHLSOLKRSFRELQVNASDGSNNKILRGCFMTPELTYTPYANSVASCERP
ERTPGELEPQTEVEHITSMAYRNRLRLRGFMTEPELTYTPYANSVASCERP
PSQPLDWPTRKRIALGSARGLSYLHRCDFRIHRDKAANTLDEFEVAGDFGL
AKLMDYKDHVTVAVRGITGLIAPETLSGKSSSEDFVGEIMLELITGQRAFDLA
RLANDDMVLDMVKGLLEKRLKEMLVDPDLOTNYEERLEDEVLVALCTQGSMPER
PKMSEVVRMLEGDGLAEKMDQKVEILREEIDLSPPNPSDWILDSTYNLHAEVLSGP
R"

BASE COUNT 569 a 419 c 502 g 599 t
ORIGIN

Query Match 51.5%; Score 934.8; DB 5; Length 2089;
Best Local Similarity 74.5%; Pred. No. 4,5e-199;
Matches 1192; Conservative 0; Mismatches 402; Indels 6; Gaps 1;

Db 175 ttgagacttacagacaataagaagccaattcctctagtgattcttgggaatcgcaca 234
Qy 486 TTGGAGCTTTACATTAACAATACATACGCGCCATTCCTGATATCTTGGAAATCTGA 545
Db 235 aattgtgagcttgagacatacatgaatagcttctctgagactataccgagacatca 294
Qy 546 AACTTAGAGATTGGATCTTTACTTAAACAGCTTCGCGGTCTTATTCGGAGATCATG 605
Db 295 ggaagacttaagaagctaaattcttgcgtctcaacaacaagaagcctctctgtccaatt 354
Qy 606 GGAAGCTTTCAAGCTGAGATTTCTCGGCTTAACACACACAGCTCATCTGGGTGAAT 665
Db 355 ccaatgtgaactatataatacaactctcaagctccttgagattataaacaacatcgcta 414
Qy 666 CCTATGTACTGACCATATTACTACCTTCMAAGTGTATATCTATCMAATAACAACATC 725
Db 415 tcaagacagtaacggatgaatgctcattcttggttacacactacagatttgccaat 474
Qy 726 TCTGTTCAGTCCCTGCAATAGCGCTCCTTCACACTTTCACACCCATCAATTTGCTAAT 785
Db 475 aattgaaattatgtgaccccgtaactcggagagcgccctgagatctccccaatttg 534
Qy 786 AACTTAGACCTATGTGACCTGTTAACAAGTCAACCAATGCTCGATGATCTCCCGTTTCT 845
Db 535 ccaccacctcgttcatccacatacaacagtaacagctccacaggaacaaatgtgtccact 594
Qy 846 CCTCCACACCTTTTATTAACAACCTCCCGAGTTCCACCCCGAGTGGATGATTAAT 905
Db 595 ggaactatgtctggggagtagctgctgctgcttcttactggttggtagcactgcaatg 654
Qy 906 GGAGCAATAGCTGTGGGTGCTGCAAGGTCTCTTGGCTTGGCTGCTGCAATA 965
Db 655 gcaattgcatgtgctggagagaagaacccgcgagagaacattcttctgtgctgcaagctgaa 714
Qy 966 GCCTTGTGCTGTGGTGGCGAGCAAGAACCCACTAATATTTTCTTCGATGTCCTGCCGAA 1025
Db 715 gagaagccaagaagtgacacttgctgaactgaagagtttctctctgagagaattgcaatgc 774
Qy 1026 GAAGATCCAGAGTTCACTCTGGACAGCTCAAGAGGTTTCTTTCGCGGAGCTACAGTG 1085
Db 775 ggaagggactttagt-----accatacttgaagaggtgagattgtgtaaggtgtat 828
Qy 1086 GCGAGTGAATGGTTTAAACAACAACATTTTGGCGAGAGGTGGTTGGGAAAGTCTAC 1145
Db 829 aagagacgcttctgtagtgcacttgtagcagttaaaggtctaaagaagaaagcaga 888
Qy 1146 AAGGACGCTTGACAGCAAGCTCTTGTCTGCTCAAGAGACTGAAGAGACGAGACT 1205

Oy 1741 tctgtccaaagataaacagcatataaattgtaagaataatttttaagtt 1797
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 6640 TCTGCTCCAAATTAACAGCATATATAATGT-AATGAAATTATTTTATGCTT 6695

 RESULT 5
 F23M19/c
 LOCUS F23M19 88401 bp DNA 17-JUN-1999
 DEFINITION Arabidopsis thaliana chromosome 1 BAC F23M19 sequence, complete
 ACCESSION AC007454
 VERSION AC007454.3 GI:5091613
 KEYWORDS HTG.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsis.
 REFERENCE 1 (bases 1 to 88401)
 AUTHORS Vysotskaya, V.S., Schwartz, J.R., Yu, G., Tortum, M., Lenz, C., Liu, S.,
 Lee, J., Liu, A., Li, J., Kremenetska, I., Luros, J., Gonzalez, A.,
 Altafi, H., Araujo, R., Chao, Q., Conn, L., Conway, A.B., Dunn, P.,
 Hansen, N., Huizar, L., Kim, C., Palm, C., Rowley, D., Shinn, P.,
 Walker, M., Davis, R.W., Ecker, J.R., Federspiel, N.A. and Theologis, A.
 Arabidopsis thaliana chromosome 1 BAC F23M19 sequence
 Unpublished (1999)
 REFERENCE 2 (bases 1 to 88401)
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAY-1999) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 3 (bases 1 to 88401)
 REFERENCE 3 (bases 1 to 88401)
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUN-1999) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 4 (bases 1 to 88401)
 REFERENCE 4 (bases 1 to 88401)
 AUTHORS Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUN-1999) Plant Gene Expression Center, 800 Buchanan
 St., Albany, CA 94710, USA
 On Jun 17, 1999 this sequence version replaced gi:4996902.
 The sequence of BAC F23M19 from Arabidopsis thaliana chromosome 1.
 FEATURES
 source
 location:Qualifiers
 1. 88401
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="F23M19"
 complement(796..3891)
 /gene="F23M19.1"
 complement(join(796..1116,1205..1271,1349..1432,
 1506..1636,1794..1850,1999..2047,2166..2251,2468..2545,
 2771..2893,2990..3079,3154..3254,3332..3393,3467..3556,
 3638..3891))
 /note="F23M19.1"
 /note="Contains a pfl00561 alpha/beta hydrolase fold
 domain."
 /codon_start=1
 /evidence=not_experimental
 /protein_id="AAD39602.1"
 /db_xref="GI:5091614"
 /translation="MEEDLILSPFDLDFQALTLPIRHYIALFVCTIFINVF
 EFRLDGAILIRFCRNLINPDSPIHGVSQCLHGRVATPWLASPHQICFL
 NFHGLPVTYRQLFLTSDGTLALDMLTNSDVLKLNKSEITKEDTPIAVVP
 GLTSDSSAYLKLADYATGANNVYLSNHRGLGVSVDCEFNAGMTDIDIVVLDY
 LQHKYRAPLPAIGTSIGANLYKYLGEEGKPLRGAVAIQSPMDLLIDRFLICRL
 KOKLYKRAITIGLOGYALHROPFLRNMGKIKRSIRDPNHNATCIYNGKEETVAF
 YKRSSSTQYGVGNVAPPLICTSALDDPLCTREARIPDECRANKRIYVATINRGGHAF
 FEGLGSSLMWVRAINFLGLVSCSPYIMHIOKIVDKRSSGSGKPEINSINOPIYLTAE

gene
 CDS
 complement(join(10781..11318,11416..11460,11670..11746,
 11875..11991,12307..12310,12443..12512,12614..12717,
 12848..12941,13174..13213,13394..13719,13800..14114))
 /gene="F23M19.3"
 /codon_start=1
 /evidence=not_experimental
 /product="F23M19.3"
 /protein_id="AAD39604.1"
 /db_xref="GI:5091616"
 /translation="MGLCSRSSSVYNNAPGCTAHVNGHLLNNASDLNHSGEGLK
 DDPSPYEVNDNKRTHSESEFPIVSSGSHSPQNTEDIPRLISRLVSKSSTYSROAA
 VAKSEVSSLLGRAGTMOGRADVDDLSLGSMMNLMISGFSATVKNKSTKLSLSE
 EVANTIVYGNALMHSLSKDSITLHKEVLLSEGVQNLISQMDLILIAADRRFQD
 PVTLYVTOBDLYHEHLADREDDYOKRIOEEEPSTAGVGDTLILHRTLEKSOAK
 HVRNKKRSLSRLIEEYMEKLVNVPVHRIHEHAFGADPDKADPADPINKKISGS
 AGLALHNTITTDITLYRSRSTPASTRDLVYGLDPSIKSALRSKISGFYKEELT
 VPQIAEMEKTLQMLVPAVNTTAHGFVGVGWSASPSANRPRGOTILRDTLH
 HADKEKTEATYILDVYVLMHLVTVQVRAITGGLSPKSPISRNQKOTIOLSSGSHNP
 SMGLPLTTEDEQEMLRDYSKRRKTPGISKSGEFTVAKARLCKHHRLSKSSSPMAG
 EMKKKKDTESTRRSPVPIIDFDIRKALVDVDRVTRSL"
 23307..26279
 /gene="F23M19.4"
 complement(join(23307..23355,23431..23627,23656..23751,23850..23906,
 24002..24145,24225..24309,24401..24491,24571..24735,
 24842..24955,25045..25117,25210..25348,25453..25865,
 25953..26140,26231..26279)
 /gene="F23M19.4"
 /note="Similar to gb|AF082176 auxin response factor 9 from
 Arabidopsis thaliana."
 /codon_start=1
 /evidence=not_experimental
 /protein_id="AAD39615.1"
 /db_xref="GI:5091627"
 /translation="MESGNYVNAQBELSGITIDGSKSYVBDMLKAGPLCIDPKLGE
 KYTFPQGHIEVLSLSLPLPESLSLHLSLSLSVSEETSTEEELNEQPIDVLDS
 KLCQVIAIHLCVKNNSDETYAETLMDTQVVPIONEQPFLNHYRGTPORHLATD
 TSAHGFEVPPKKHIAECLPSLDMSQPLPAQCLILIDHGNMRNHHYRGTPORHLATD
 TGNMFTSKKLVAGDIVFVRGEGELRGVCIARRHQGNISSIVSICMRHGVYA
 SAKHAFDNOCMEYTVYKPRSSKFLVSYKQFLFADYANNKFNSSRTFMLEGGDESERG
 EGTITGVSDDESPHMKCSFMRSLFQWDEFTFSPKPKVSPMDIDHMLPAINPVSFL
 NKKRLREYNEIGSSSHLPLPLTLOGQENELSVASPNKISLRTRDTEDMANPSKLL
 MSYPQMPKLYNNQWTEMEENTTTGTGTFNLFVGLTDPVILDPVILDEIGSEIS
 KLTEBKKFGLSQTLSRPTQIONKOFSSSRCTGKLVQMGVTLGRAVDLSVNGDQLIL
 ELKEFLDIKGOJOTRNQWEIAFTDSEDKMLVGDPPWPEFCNMVKKIFIQRR"
 complement(28888..31377)
 /gene="F23M19.5"
 complement(28888..31377)
 /gene="F23M19.5"
 /note="Contains similarity to gi|479356 protein kinase PK1
 from Zea mays, is a member of the pfl00954 S-locus
 glycoprotein family and contains a pfl00069 Eukaryotic
 protein kinase domain."
 /codon_start=1

Query Match	23.9%	Score 434	DB 8	Length 88401
Best Local Similarity	64.2%	Pred. No. 4,36-87		
Matches 828	Conservative 0	Mismatches 245	Indels 217	Gaps 3
697	tttgaatgcccacgtctaaagagaccacgaagatgacccctgtgtcaactgaagagagtttct	756		
Db 73350	tttctgcatatcaccacacgaacacccctttagcttcttcttgggacgactttttagcggttctc	73291		
757	ctcgagaaattgcaagtcgcaacgagatctttag-----taccactactggaagagt	810		

Dh	73230	CTAGGGAACTTCATACATACCACTGATACGCTTCAGCAACAAGAAACATTTGGGCGAGGT	73231
Qy	811	ggatttgtaagtgataaaggagcctctcgaatgctcacttgacgaattaaag	870
Dh	73230	GGGTTCCGAAAGACTCTACAAAGCCGCTTCTCATGGAACACTTGTGTGACCAACGG	73171
Qy	871	cttaagaagaagaccacgaagtgtagtgactgaattcaacaagaaggtggaaatgatt	930
Dh	73170	CTTAAAGAAAGGAAACCCCAAGTGGCACTCCAGTTTCAGACAGAGTGGAGATGATA	73111
Qy	931	agcatgctctgcatcgaatactctctgactagtggtttctctgcatgacacacaagaag	990
Dh	73110	AGCATGCGCGTTACAGAAATCTCCTACGCGTACGCGGTTTCTGTATGACCCCTACCGAG	73051
Qy	991	cgcgtctctgataatccaacttgctaatggaagtgttcgctcgttgattaaag	1044
Dh	73050	AGATGGCTTGTATCTTACATAGGCTAATGGAAGTGTCCGCTTCCTTTGGAAGGTAAAC	72991
Qy	1044	-----	1044
Dh	72990	CTTGGAATTTAACTGTTGTATCATAAGTAAAGAAAGACGCCCAATGATGATAGAG	72931
Qy	1044	-----	1044
Dh	72930	TTGTTTTCATCTTATCCATTTTAAAACTTCCAAATTCATTCAGTGAACCTTCTTGAA	72871
Qy	1044	-----	1095
Dh	72870	ATGATTACAGAAAGCTCCACCATCAGTTGCTCTACGCTGGTAAATAAACAGCAAAAC	72811
Qy	1096	gcactaagatctgctagggggtcttctattctgcatgacatctgtaaccagaattatc	1155
Dh	72810	GCCTAGGAATACAGAGGGGTTGCTTATCTTCATGATCATTCGACCCCAAAATTAAT	72751
Qy	1156	catgttatgtaaagctgcaatatattattgagcgaagaattgggctgtgtgaagt	1215
Dh	72750	CACCGTATGTGAAGTGTCTAATATCTGTGTGACGAGGAATTTGAGGGGTGATAGGT	72691
Qy	1216	gatttgggttagctaggtctcatgataccaagaataccatgttaacaactgctgaag	1275
Dh	72690	GATTTCGGGTAGCTAGACTTAAGACTTAAGACTTAAAGTACTCATGTACACAAGCGCTGGCGT	72631
Qy	1276	ggtaccttgggtctacatagctcccgagtaacctctctgactgtyaaagtcataagaagaacc	1335
Dh	72630	GGGACTATTGGAACAATTGCTCTGTGATCTCTCACTCGAATAATTTCCAGAGAAACT	72571
Qy	1336	gattcttctggtatggatattgctccttagagctcatctacttgacagagaagctttgat	1395
Dh	72570	GATGTTTTTGCTACGGGATCAAGCTTTTGGAACTGATTAAAGTCAAGAGCTTTTGAT	72511
Qy	1396	cttgcctgccttgcgaacgaatgatagttattgtttgttgaatg	1440
Dh	72510	CTTGCAAGACTGGGCAATGACGATGACGTTATGCTCTAGAAATTGGGTAAACACAGATCC	72451
Qy	1440	-----	1440
Dh	72450	TTTAAACAATCTGCTACTCTCTCAAAAAGCTGAATTAATCTGTTCAATTTGGCTTCTCA	72391
Qy	1440	ggttaaaagccttttgaagaagaanaagttgtagatgtctgtgcatcctgaacctgtagaa	1499
Dh	72390	GGTGAAGGCTTTTGAAGGAAGAAGAGCTGGAGATGCTTGGATGCTCTGACCTGCAGAAAG	72331
Qy	1500	caattacattgacacagaagtctgagcagcttatccaagtacatctctgtlaccagaag	1559
Dh	72330	CATTTTACACAAACCAAACTAATAACAGCTCAATAAAGTGGCTCTTCTTCGACACACAGAG	72271
Qy	1560	ttcgcgaatgtagggccttaagaatgtccaagaatgaccgaatgcttgaagttatggcct	1619
Dh	72270	CTCACTTATGGAAGACACTTAAGATGTCTAGGTTGTGTGATTCCTGAAGTGAACGTTT	72211
Qy	1620	tgcagaanaagtggagcagatgtgcanaaagttagaagtcacatcaagaacgttagaatgac	1679
Dh	72210	AGCGAGAAAGGGGACACAGTGGCAGAAAGTGGAAAGTCTTCAGGCAAGAAAGTGGAGCTTTC	72151

```
Oy 1680 tccacatcgaactctgcatgacatcgaactcgaacataactgcatgattgaatt 1739
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 72150 TTCTACACCCACCTCTGATCGATCTGATTCGACATATACTTCAATGAGACT 72091
Oy 1740 atctggtccaaagataaagacataataat 1769
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 72090 GTCTGTCACAGATAACGACATCTGTAATT 72061

RESULT 6
AFF17M5/c 96475 bp DNA PLN 11-MAR-1999
LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone F17M5 (ESSA
DEFINITION project).
ACCESSION AT035678
VERSION AT035678.1 GI:4490291
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
Rosidae; Capparales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 96475)
Bevan, M., Rose, M., Hempel, S., Entlan, K., -D., Hoheisel, J.,
Mewes, H.W., Mannhaupt, G., Mayer, K.F.X. and Schueeller, C.
Unpublished
2 (bases 1 to 96475)
EU Arabidopsis sequencing project.
REFERENCE
AUTHORS
DIRECT SUBMISSION
TITLE
Submitted (11-MAR-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
schueeller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@brc.ac.uk
COMMENT
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosome 4 can be
viewed at: http://webserver.mips.biochem.mpg.de/proj/thal/.
FEATURES
source
1. 96475
/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
/377..3928
/gene="F17M5.10"
/complement(2377..2571)
/number=1
/complement(join(2377..2571,2755..2837,3032..3098,
3180..3230,3461..3574,3668..3724,3815..3928))
/Note="similarity to Caenorhabditis elegans cosmId, gene
T1661.11, PID:el349366; contains EST gb:U20620, N5143,
T45799, H36819, R65238, N37318"
/codon_start=1
/product="putative protein"
/protein_id="CAB38783.1"
/db_xref="GI:4490292"
/translation="MGVEIQSPOSSYTVROIALNPENPILPDLENYVNTSORY
SLVENICLRXQPEPEMNMHIYARILYKALAMPDPDSICFILPERYQEEQK
SLIYSHLETGRQOQFWEAKRNKHILEAVPGEOAIQATASHLSLSTOKPRSVL
AEAVMDGASLDKFLIQQVTSNGMIVEEGSVLPQNEFNHPELKRKNGENPVEHI
ARIPFLLG"
/complement(2572..2754)
/number=1
/complement(2755..2837)
/number=2
/complement(2838..3031)
/number=2
/complement(3032..3098)
/number=3
/complement(3099..3179)
/number=3

exon complement(3180..3230)
/number=4
/complement(3231..3460)
/number=4
/complement(3461..3574)
/number=5
/complement(3575..3667)
/number=5
/complement(3668..3724)
/number=6
/complement(3725..3814)
/number=6
/complement(3815..3928)
/number=7
/4544..6491
/gene="F17M5.20"
/complement(4544..4687)
/number=1
/complement(join(4544..4687,4781..5362,5443..5712,
5815..6093,6423..6491))
/Note="strong similarity to WD-repeat protein -Daucus
carota, PID:g2253631; Contains Trp-Asp (WD-40) repeats
signature [VLAIALDHTVYLMDA] [TGGGGGDRTIKFWMT]
[VASAAGDETLRFWMV]"
/codon_start=1
/product="WD-repeat protein-like protein"
/protein_id="CAB38784.1"
/db_xref="GI:4490293"
/translation="WDAGLNRCPIQEHFLPKRKNENIDRFIPRSMANPDYAFALT
ERRKRDQSAIVSPSKRAYRQIAETNMLNHTRLIARNRPOAPVELLPNSASH
QOPKSVPRPRYIPQTSERTLADPIDVDIFYNLIDWGSANVLAIALDHTVYLMDQSY
SYSELVYTIDEEGPPTSINMAPDGRHVAVGLNNSVQIOMDSASROALTILKGQSHV
GSLANMHLITFGMDGLIINNDVIRSPVETRYGHTQECGLKWSGGQIASGNN
DNVAHIMNRSVASSNSTOWMLHREHHSAYKALAMCPFOANLLATGGGGDRITKFM
NHTGACANSVDTSQVCSLMSKRRRLISHQFTONQTLNMTKPSVYKAAELTGHT
SKVLLMAQSPDCTVASAAGDETLRFWMVFGVPEYAKAAKAAVAEPESHVNRIR"
/complement(4688..4780)
/number=1
/complement(4781..5362)
/number=2
/complement(5363..5442)
/number=2
/complement(5443..5712)
/number=3
/complement(5713..5814)
/number=3
/complement(5815..6093)
/number=4
/complement(6094..6422)
/number=4
/complement(6423..6491)
/number=5
/7856..9901
/gene="F17M5.30"
/complement(7856..7999)
/number=1
/complement(join(7856..7999,8104..8685,8767..9036,
9139..9417,9803..9901))
/Note="strong similarity to WD-repeat protein -Daucus
carota, PID:g2253631; Contains Trp-Asp (WD-40) repeats
signature [VLAIALDHTVYLMDA] [TGGGGGDRTIKFWMT]
[VASAAGDETLRFWMV]"
/codon_start=1
/product="WD-repeat protein-like protein"
/protein_id="CAB38785.1"
/db_xref="GI:4490294"
/translation="WDAGNNNTSSHYKTQARCPQIHFLLPKRKNENIDRFIPRSM
NPDYAFALTREGKRGDQTAAVSPSKRAYRQIAETNMLNHTRLIARNRPOAPVEL
LPNSASHQOPKSVPRPRYIPQTSERTLADPIDVDIFYNLIDWGSANVLAIALDH
TYLMDASGTSYSELVYTIDEEGPPTSINMAPDGRHVAVGLNNSVQIOMDSASNNQIR
TIKGGHOSRVGSLANMHLITFGMDGLIINNDVIRSPVETRYGHTQECGLKWSG
SGQQLASGNDNVYHIMDRSVASSNSTOWMLHREHHSAYKALAMCPFOANLLATGCG
GGDRTIKFWMTHTGACINSVDTSQVCSLMSKRRRLISHQFTONQTLNMTKPSV
```


ACCESSION AC006436 AE002093
 KEYWORDS AC006436.4 GI:6598565
 SOURCE HRC.
 ORGANISM thale cress.
 Arabidopsis thaliana
 Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicotyledons; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsids.
 REFERENCE 1 (bases 1 to 65899)
 AUTHORS Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D.,
 Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E.,
 Feldhym, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Ronning, C.M.,
 Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., VanAken, S.E., Umayam, L.,
 Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H.,
 Goodman, H.M., Somerville, C.R., Cohenbayer, G.P., Preuss, D.,
 Nleman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and
 Venter, J.C.
 TITLE Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana
 JOURNAL Nature 402 (6763), 761-768 (1999)
 MEDLINE 20083487
 REFERENCE 2 (bases 1 to 65899)
 AUTHORS Lin, X.
 TITLE Direct Submission
 JOURNAL Submitted (13-DEC-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 COMMENT On Dec 17, 1999 this sequence version replaced gi:4726109.
 The sequence and annotation of chromosome 2 were merged from those
 of the individual clones on this chromosome after removing
 overlaps. For detailed information, please see the TIGR web site
 (<http://www.tigr.org/tdb/at.html>).

Genes were identified by a combination of three methods: Gene
 prediction programs including GRAIL,
 (<http://archur.epm.ornl.gov/pub/xyrall>), Genefinder (Phil Green,
 University of Washington), Genscan (Chris Burge,
<http://genomic.stanford.edu/GENSCANW.html>), and NetPlantGene
 (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the
 complete sequence against a peptide database and plant EST
 databases at TIGR, and manual curations based on those analyses.
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as 'unknown' proteins.
 Genes without protein or EST similarity, that are predicted by two
 or more gene prediction programs over most of their length are
 annotated as 'hypothetical' proteins. Genes encoding TRNAS are
 predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were
 identified by repeatmasker (Arian Smil,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are
 numbered from the top to the bottom of the chromosome.

We thank the CSHL/Mashu/ABI consortium for sequencing BAC clones
 F6P23, F5J6, T17A5, and T13J16, the ESSA group for sequencing clone
 F13J4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards
 and Satoshi Tabata for helpful assistance. In addition, we would
 like to thank the TIGR Bioinformatics Department, especially Lixin
 Zhou, Hanif Khalek, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy
 Peterson, Michael Holmes, and Delwood Richardson for software and
 database support.

This work was supported by the National Science Foundation,
 Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.
 Location/Qualifiers
 1. 65899
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="2"

misc_feature 1..65698
 /note="Sequence from clone F13J11"
 /complement(join(451..842,907..11234))
 /gene="At2g13660"
 gene complement(451..1234)
 /gene="At2g13660"
 /note="F13J11.1"
 CDS complement(join(517..842,907..1114))
 /gene="At2g13660"
 /note="unknown protein"
 /codon_start=1
 /protein_id="AAD28310.1"
 /db_xref="GI:4726110"
 /translation="MOCGMREKGFAYDSSSPSTFISCHRLITGNFLHPADOL
 SPNNYLICRLTYLALACHRSPEIODEDSNSYSVPIINRLINLPRTALANKPF
 WRKSELDSIGRIRIHSNNKDPDFNNFHLPEIGRYVNLISQRAIKRKKNVKNSEK
 ETGIKTNLHKPKSSD"
 complement(2606..2687)
 /rpt_family="GAA)n"
 join(<3349..3541,3644..3744,4109..4230,4375..4471,
 4583..4693,4792..4928,5629..5755,5848..5997)
 /gene="At2g13670"
 <3349..5997
 /gene="At2g13670"
 /note="F13J11.2: predicted by genscan and genefinder"
 join(3349..3541,3644..3744,4109..4230,4375..4471,
 4583..4693,4792..4928,5629..5755,5848..5997)
 /gene="At2g13670"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="AAD28311.1"
 /db_xref="GI:4726111"
 /translation="MAOSTSHDPOGLARRPSRAATVTSIEVDHEVNASLGTI
 APIIRVAETIEREPYAVYLCRFAPFAKARLDPPSSGSGRGVQFKTLQORERDAS
 SLARKYKDGREVSFPOOYEVYVALDGOADRAQGRAYOTAGVLEFMAVN
 KSEYEVAPEPITIAADVDEKNEIYTPYIILDSGASQSMQEEYKAAVALGN
 TRGLNWSGFEQHRKKTGNDLDMWAMFGFQASWPCNAQDIDQKRLIMGLYLI
 WGEANIREMECCYIFHMAVEYELHGLAGNVSIVYGENIKPSYGGDEAPLAKVIT
 PIYRVQVVOV"
 10367..13711
 /gene="At2g13680"
 /note="F13J11.3"
 /pseudo
 complement(join(<14929..15843,16401..>17114))
 /gene="At2g13690"
 complement(<14929..>17114)
 /gene="At2g13690"
 /note="F13J11.4: predicted by genscan and genefinder"
 complement(join(14929..15843,16401..17114))
 /gene="At2g13690"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="AAD28312.1"
 /db_xref="GI:4726112"
 /translation="MGDSDLAKPNNFCTGHLRSRWSGCFAPAPSPDRSSRNH
 IPAKQOORAKVCSPOSSKSAINYNRDLPRILSPGVSTIDSPYTTMQEET
 TQEEEDDAVVDSPNLSKESFRAPKLEVGSGISSEYDARSLKGRNGGVAVLELS
 LEVLAANDVSGLIAEKKKSSSSSLGKNCRIKIVCVENMGVREVEVLEPES
 NVLIKRFMTGVYRAIDVLEVAAGIKRSRAVLSCKLYLEAVPTDEEEKLRLLGIY
 SFDDAVAEIILARNSETNEDLOSLKLYWSTGSDVPNRELKSLVKGILCKSS
 VYERKQPIINKEIDYRAGKCCVDSLAKFEGSSSSSKREKRLISISEVENIML
 LEIMIDRIAEFPEYINGKORLYEMERKSPRYRYRYTAIRFIMAKRKRYQCG
 EAKAGVEAMFRPMIVDFGMLOKCKGKGLDMREVEBGQTLLTLPVKEQOYVMEWR
 WFSKRGTCPNLSKAFQIMWRSFLRGVESSSTOR"
 complement(18082..20371)
 /gene="At2g13700"
 /note="F13J11.5"
 /pseudo
 23450..29498
 /rpt_family="POLY_A"
 complement(29932..30854)
 /gene="At2g13710"
 /note="F13J11.6"

AUTHORS Annen, F. and Stockhaus, J.
TITLE SDRK1, a receptor-like protein kinase of Sorghum bicolor (L.)
JOURNAL Moench that is expressed in mesophyll cells
MEDLINE Planta 208 (3), 420-425 (1999)
REFERENCE 99312726
AUTHORS 2 (bases 1 to 2278)
TITLE Stockhaus, J.
JOURNAL Direct Submission
Submitted (14-AUG-1997) J. Stockhaus, Institut fuer Entwicklungs-
und Molekularbiologie der Pflanzen, Heinrich-Heine-Universitaet,
Universitaetstrasse 1, 40225 Duesseldorf, FRG
FEATURES
source
1. .2278
/organism="Sorghum bicolor"
/cultivar="TX 430"
/db_xref="taxon:4558"
/lab_host="XLI-Blue"
/tissue_type="leaf"
/clone="CRK1"
164. .235
/gene="RLK1"
164. .2023
/gene="RLK1"
/note="accumulates in mesophyll cells"
/codon_start=1
/product="putative protein serine/threonine kinase"
/protein_id="CAB51480.1"
/db_xref="GI:5596996"
/translation="MRHSEIKLIAFYILMGCOSSSLCAVDSQVAVLYEMKOLY
DNKGLSDMDNDSPCYWEIYNCDNKKVSTITSSSGITGLSLINKLTTLQDLK
DNNITGIGLIPERGNLSLTILNLGNNGSIPLSLQLSLQLDLSHNLGNIPL
SFSNPSPNDIMLAYNNSIGBIPOHLQAAHYNTGHNLCNGONLFCPEGSTMTGG
SKNSKLVIGSIGAVLCTVALVILWQMRKRPETFDIVSGONDMHEFQIKR
FSEKREIATNYSFSEONVLKSGFEGYGVYGPDSIKIAYKRLFNVERHEGLATL
REVELISIAVKNILRLIGECTPPERLILYPMENCLVASLRDRIKTEYVLDMSR
MRALGASGLIETIHERCNPKIINHDKAANILDSNENAVYDGLKMMIDIGNY
TIGRGTMGHIAPEYIKIGRPSYKTDIIGTGLMLSEIYGERALIFHEDRMEGEIM
LIDQVLMMEGRLLDLDVDRNGVYNLEELKVQIALCTHMPESRPTMSEYVOM
LKEGIYPAERWEMOIALELDROROHENMOGKLFNFSERSLINIOEAIELSTGR"
gene
164. .2023
/gene="RLK1"
233. .2020
mat_peptide
/gene="RLK1"
/product="putative protein serine/threonine kinase"
BASE COUNT 690 a 445 c 530 g 613 t
ORIGIN
Query Match 17.5% Score 317.6; DB 8; Length 2278;
Best Local Similarity 54.8%; Pred. No. 4.9e-61;
Matches 826; Conservative 0; Mismatches 634; Indels 48; Gaps 8;
QY 197 taagtggaccaatctcctagtagctctgtggaaatctgacaatcttggtgagcttggaacctat 256
DB 411 TTAACGAGAACCTTATCACCCGACCATTCGTAACATAACATTATACAGCACTGAACCTGG 470
QY 257 acatgaaagttctctgagacctatccggaacacattagaagaagcttacaagcttaagt 316
DB 471 ATACCAACAAATACATGGAAGGATCCCTCCGAGATTGGAACCTATCAAGTTTGACAA 530
QY 317 tcttgagctcaacaacacagccctctctgccaatctccaatctgaactgaactaatat 376
DB 531 TTCTTAACCTTGGAAAGAAATTAATTAAATGATGATACCAAGACTCCCTTGGACAACCTT 590
QY 377 caaccttcaagctcctgattatcaacaacatcgctacagaacccagtcacgacaag 436
DB 551 CTAAACTCCAAATTCGATCTGAAGTCACATCTTAATAGTGGGAATATCCCAAGCTT 650
QY 437 gctcaatttct---ttggttacacatacagtttgccaataattgcaattatgtgga- 493
DB 651 TCTCAAACTCTGCATCATTTGAATGATATATCTAGCATTAATAATATTAAGGGGTGAAA 710
QY 493 -ccgtaactggagagccctcgcttgatctcccaatttgcgacacacactcctgcat 551

DB 711 TCCACACACTCTACTTCAGGCGCTCATTAACAATTACAGGCAACCACTTGAAATGTC 770
QY 552 ccacacatacacagt-----aaagctccaggaacaaatgtgccca 592
DB 771 GCCAAATTTATTTCCATGTGAAGAGGAGACAAATATACAGGTGGATCGAANAATTCGA 830
QY 593 ctgagcatctctgaggaggaagtaagctgtgtgtcgtcttactctgttgcgtcacctgca 652
DB 831 AGCTAAAGGGTAAATGGAACCAATTCGAGAGCTGACCTTGTGTTACTACTACTCTC 890
QY 653 tggcatttgatgtgtgagagaa---gaaacccgagaaacatttcttattgtgcaag 709
DB 891 TAGTTTGCATGTGTGCAAGATGCGCTATCGCGCTCGAATCTCTATTAAGCTGTACG 950
QY 710 ctgaagggagcccaagaagtgcacctgtgtcaactgaagaagtttctctcgagaatgac 769
DB 951 GTCAGATGATCATATGCTTGTGAGTTGGCAATTAACGGGTCTCATAGCGAGAGCTTC 1010
QY 770 aagtcgaacaggaacttactagt-----accatacttggaagagtgattgtgaag 823
DB 1011 AGATTGCACCAACTATTTTCAGTGAACAAATGTTCTTGCAAGGCTGTTTGGTAAAG 1070
QY 824 tgaataaggagcctctgtgtgatggtcacatgta-----gcagttaaaaggtctaaag 877
DB 1071 TGATTAAGGAGATTCTTCAGGTCAGACAGATATAGATTCAGATGAACGACTATTCA 1130
QY 878 aagaagaacacacaggtgtgtgagctgcagtttcaacaagagtggaatgtattgttcag 937
DB 1131 ATGGGAGAGCGCATGAAGGGGAATTTGGCTTCTTGAGAAATTAATGATTAAGCATTG 1190
QY 938 ctggtgacgcgaatctctctgctcactagctgtgttctgtcatgacacacaagaagcgcttc 997
DB 1191 CTGTCACAGCAAGAACATATTAAGTGTATGATGATGTGTCACACCAACAGAGAGCTCT 1250
QY 998 ttgatactcatcatcagctaaagtaagtggtgtgtgtgttgaagaagcgctacgcat 1057
DB 1251 TGGCTAACCTTTCAATGAGAAATCTTAGTGTCTTCGTTTGAAGATTAATAATTAA 1310
QY 1058 cagaagctccctctgatttgccaactagaagaagagatgcacataagatctgataagggc 1117
DB 1311 ATGAAACAGATTAATGATTTGGTCTACAAAGATGGAATGCTTGTGTGCTGCGGTGTT 1370
QY 1118 ttcttatcttgcatgaccattgtgtatcccaagaattatccatcgtgtatgtaaagctgaa 1177
DB 1371 TGGATATCTTTCAATGAGCACTGCAATCCCAAGATCATCCACGCTGACGTCAAGGCTGCA 1430
QY 1178 atatattatggacgaagaattgagagctgtgtgaagttgtattgttaagctagctca 1237
DB 1431 ATGTCTTGCTTGATGGAACTTGCAGACAGTGTGAGGGAATTTGGGTTAAGCCAAAGATGA 1490
QY 1238 tgaattacaaggaataccatggtacacatgctgtgtaagggtacacttgggtctacatagctc 1297
DB 1491 TGGCAATAGGAGAGAAATACGTAACCAACAGGGGTTGTGGGCTATGGGCCACATAGCTC 1550
QY 1298 ccgagatgacctctgagactggagaagctatcagagaagaacgaattcttcttgatttgagata 1357
DB 1551 CTGAGTACATTAAGACAGGAAGGCCATCACTGTAAGACAAATATCTATGATATGTGTGCA 1610
QY 1358 tgccttagagctcatctactgtgacagagactttg-----atcttgctcgcttgca 1411
DB 1611 TGTGTTAGAGATTGTACAGGAGAGCTGACATTCATTCATCTGACCTGACCTATGAGAG 1670
QY 1412 acgatgtatgttatgtttgttgattgtgttaagaccttttgaagaagaaagtgtgg 1471
DB 1671 AAGCAAGGGAGATATGCTATTTGATCAAGTCAAGCTATGATGAGAGGAAGGCGCATAC 1730
QY 1472 agatgtgtgcatcctgtgacctgtgagaacaatbatatgacacaagaagtgtgagactta 1531
DB 1731 TCGACCTAGTAGATCGCAATCTAGGTGGCGCTACACACTTGAAGAGCTGAGAGGTTA 1790
QY 1532 ttcaagtagcatctactctgtacccaggggttcgccaatgtgaagcggtcctaagaatgtcaagg 1591

Db 1791 CCAGATAGACATCTCTGACCCACATGGAACCTAGTCAACGCCCTACATGTCAGAG 1850
Qy 1592 tagtccgaatcgttgaagatga---tggccttcgagaagaatggagagtgcaaaag 1648
Db 1851 TCGGAGATGCTGGAAGAGAGATTTCTCCACAGAGAGGCTTGAGAGATGCACTCG 1910
Qy 1649 tagaagtc 1656
Db 1911 CTGAGCTC 1918

RESULT 10
LOCUS A67815 4081 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 20 from Patent WO9743427.
ACCESSION A67815
VERSION A67815.1 GI:4756638
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE 1 (bases 1 to 4081)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICYTIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source location/Qualifiers
1..4081
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="ARABIDOPSIS SEBK GENE"
exon 1280..1367
exon 1796..1928
exon 2014..2085
exon 2203..2346
exon 2450..2521
exon 2617..2688
exon 2772..2884
exon 3015..3146
exon 3305..3646
exon 3760..4081
BASE COUNT 1120 a 770 c 785 g 1406 t
ORIGIN

Query Match 14.9%; Score 270.8; DB 5; Length 4081;
Best Local Similarity 64.6%; Pred. No. 1.4e-50;
Matches 504; Conservative 0; Mismatches 157; Indels 119; Gaps 2;

Qy 707 cagctgaagaggaccgaagtgactgtgtaactgaagaggtttctctcgaggaat 766
Db 3302 CAGCGGAAGAGATCCAGAGATTCATCTGGGACGCTCAAGAGGTTTCTTCGGGAGC 3361
Qy 767 tgcagtcgaacgatacttagt-----accatacttgaagagtgatgttga 820
Db 3362 TACAAGTGGCAGTGAAGTGGTTAGTAAACAAGACATTTGGGAGAGGTTGGCA 3421
Qy 821 aggtgtataagggagcgcctgtctgtgctcactgtgacgaagttaaaagcttaagaag 880
Db 3422 AAGCTTACAAAGGAGCGCTTGGCAGACGAACTCTTGTGCTCAAGAGACTGAAGAG 3481
Qy 881 aacgaacacaggtgtgtagctgtgcaattcaaacagaggtggaatattagcagctg 940
Db 3482 AGCGAATCCAGGTGGAGAGCTCCAGTTTCAACAGAAAGTGAATATTAAGTATGGCAG 3541
Qy 941 tgcacgaacatcttcgtctacgtgtgttctgcatgaacacacagacggtctcttg 1000
Db 3542 TTCATCGAATCTGTGTGAGATTAGAGGTTCTGTATGACACCGAGAGAGATGCTTG 3601
Qy 1001 tatatccatacatggtcgaatgaggtgtgtgtgtttaaag----- 1044

Db 3602 TGTATCCTTACATGCGCAATGGAAGTGTCTTGTGTCTCAGAGGTAATAAATAACAA 3661
Qy 1044 ----- 1044
Db 3662 TTAACATCTTGTCTCTCTCAATTAATTGACGTGAAGTGTTTTTCATGTTTCT 3721
Qy 1044 ----- agagcgtcagcaatcaaacctcc 1067
Db 3722 TTATGGGTTCAATATGTTGGTTACACTAATGACACAGAGAGGCCACCGCTCAACCTCC 3781
Qy 1068 ccttgattggccacttaggaagagtagtgactgaagatcgtcagggggtcttctatt 1127
Db 3782 GCTTGATTTGGCCACGCGGAGAGATCGCGCTAGGCTGAGCTCGAGGTTGCTTACCT 3841
Qy 1128 gcatgacatgtgatcccaagattatcactgtagtgaataaagctgcaaatattatt 1187
Db 3842 ACATGATCAGTCCGATCCGAAGATCATTCACCGTGAAGTAAAGCAACATCTCTT 3901
Qy 1188 ggaagaaatttgaggtgtgttgatgttggttgatgtagctgaagctcattaca 1247
Db 3902 AGACGAGAAATTCGAACCGGTTGTGAGATTTCGGGTTGGCAAACTAATGAGACTATA 3961
Qy 1248 ggaaccatgttaactcgtgaaggggtacctgtggtgactatagctccagtaact 1307
Db 3962 AGACATCTACCTGACACACAGAGTCCGTGGCACCATCGTCAATGCTCCAAATATCT 4021
Qy 1308 ctgcagtcgaagatcagaagagaccgagatgtcttggttatgtagtattgctctaga 1367
Db 4022 CTCACCGGAAATCTTCAGAGAAACCGAGTTTCGGATGAGATCATGCTCTAGA 4081

RESULT 11
LOCUS AB013395/c 86064 bp DNA PLN 20-NOV-1999
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MON23,
complete sequence.
ACCESSION AB013395
VERSION AB013395.1 GI:3128142
KEYWORDS HMG.
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone:lib:Mitsui P1
clone:MON23.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE 1 (sites)
AUTHORS Nakamura,Y.
TITLE Structural Analysis of Arabidopsis thaliana Chromosome 5. VI
JOURNAL Unpublished (1998)
REFERENCE 2 (bases 1 to 86064)
AUTHORS Nakamura,Y.
TITLE Direct Submission
SUBMITTED (06-MAY-1998) to the DDBJ/EMBL/GenBank databases.
Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of
Gene Structure 2; 1532-3, Yana, Kisarazu, Chiba 292, Japan
(E-mail:ynakamuekazusa.or.jp, Tel:+81-438-52-3935,
Fax:+81-438-52-3934)
FEATURES
source location/Qualifiers
1..86064
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MON23"
/clone="lib:Mitsui P1"
BASE COUNT 27710 a 15313 c 14986 g 28055 t
ORIGIN

Query Match 13.9%; Score 252.6; DB 7; Length 86064;
Best Local Similarity 54.9%; Pred. No. 1.6e-46;

Matches 675; Conservative 0; Mismatches 404; Indels 150; Gaps 3;

```

QY 707 cagctgaagagagaccgaagtgacacctgtgtcaactgaagaaggtttctctgcgaat 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74141 CAGGTGAAGTGGACAGAAAGATGCTTTGACAGAGTTGAGAAAGTTGCATGGAGAGC 74082

QY 767 tgcagctgcgaacggatactttagt-----accatacttgaaaggtgatttgta 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74081 TTCAGTGGCTACAGATGAGTGAAGAAAGATGTTCTCGCAAGAGAGGCTTTGGCA 74022

QY 821 aggtgtataaaggagacctgtgctgactgtgacgttaaaaggcttaaaag 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74021 AAGTTTACAAAGATGCTTTGCGATGACCAAGATCGCTGTAAGAAAGTTGACGATT 73962

QY 881 aacgaacacccaggtggtgagctgcagttcaacagaagtgaaattagatgagctg 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73961 TTGAAGCTGCACAGAGAGATGATGATCCAGAGAGAAATGAGATGATTAAGTACTG 73902

QY 941 tgcaggaatctctgtgctgaagtggttctgcacagaccaaagagagcttctg 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73901 TTCAATAGGATGCTGCTTATCGCTTTGACACCAAAAGTGAACGACTTTGG 73842

QY 1001 tatatccatcacatgctgaagtggtgctgctgttaag----- 1044
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73841 TGTATCTTTTCATGAGATCTAAGTGTTCGATATTGCTTAAGAGGTAAAAAGAGACA 73782

QY 1044 -----agag 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73781 TAAGCTATTAACCTTTGATTTTCAGAGTTTAAGATTTTGATCTATATCTATGCGAG 73722

QY 1048 cgtcagccatcacgaacctcccttattgtgccaactgaagaagattgacatgactc 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73721 ATTTAAACCGGGGAGCCAGTTCTGATGTGTCAGAGAGAAACAGATTGCTTATGAGCA 73662

QY 1108 gctaggggcttcttattgtcagatgcatgcatgcatccaaagattatccatgtagta 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73661 GCACAGAGCTCGAATATCTTCATTAATGCAACCGAAGATCAACACAGAGATGTG 73602

QY 1168 aagctgcacaatatatttgacgaagaattgagagctgtgtgaagtgattgggtta 1227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73601 AAGCTGCCAAATGTTTACTAGATTAACAACCTTGACAGAGGTGTGATTTGGTTTA 73542

QY 1228 gctaggtcatgattacaagatacccatgttaacactgctgaaggggtacctggc 1287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73541 GCCAAGTTGGTATGATAGAGAGCTATGATACCACTCAGTGCCAGAGAACATGGGT 73482

QY 1288 tacatagctcccgagtaacctctgcactggaagttcatagagaagacgagatgcttgc 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73481 CATATTGCCACGAATGTATATCCACAGGAAATGCTAGGAAACCGATGTTTCGGG 73422

QY 1348 tatggatatagctcttgaagctcatctgacagagagactttgattctgtcgcct 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73421 TACGGATATATGCTTCTGAGCTTGTAACTGACAAAGAGCAATGATTCTCGGGTTA 73362

QY 1408 ggcgaacgtagatgattatgctgtga----- 1437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73361 GAGGAGAGAGATGATGCTTATGCTAAGACATGTAAGTAAATGTTGATCTTTTGCTTC 73302

QY 1437 -----ttgggttaaaagcctttga 1456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73301 TTTTCAAGCTTCTCATATATGTTAAATTAACAATCTTAGTTCAGGTGAGAAATGGAAA 73242

QY 1457 aagagaaaaagtttgatagatgctgcgtcgaactcgtgacgtgaaagaacattacattga 1516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73241 GAGAGAGAGATTAGAGACATAGTAGTAAGAAAGCTTGATGAGATTATTAAGAGAG 73182

QY 1517 aagttgaagacttatacagtagactactgtacaccaggttcgccaatgtagcgcgc 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73181 AAGTTGAATATGATAGATAGAGTAGCTGTGATGACACAAAGAGCACCAGGAAAGAGAC 73122

QY 1577 ctaagatgtcagaagtagtccgaatgcttgaaagtgcgtgcctgcgaagaaagtggagc 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73121 CAGCGATGTGGAAGTAGTAGAATGCTAGAAAGGAGAGGCTTGCAAGAGATGGGAGAG 73062

```

```

QY 1637 agtgcacaaagtagaagtcattccatcaagacgtagaattagctccacatcgactctg 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73061 AGTGGCAAAATCTTGAGTGTGACGAGCAAGAGATTTCAGAGTTGACAGAGAGTTTG 73002

QY 1697 aatgatacttagactgcacagataacttgcatgctttaaattatctgttccaagataa 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73001 ATTGGGAGAGATTCATTAATATCAATCAAGATGCTATTGATTAATCTGGTGAAGTAGA 72942

QY 1757 cagcatataaagtgaatgaattata 1785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72941 AACAAAAATATCTTTGATTCAAAAGATA 72913

```

```

RESULT 12
AB025639/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

Arabidopsis thaliana (strain:Columbia) DNA, clone_11b:Mitsui P1
clone:MML2.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsids.
1 (sites)
Nakamura,Y.
Structural Analysis of Arabidopsis thaliana Chromosome 3. II
Unpublished (1999)
2 (bases 1 to 84896)
Nakamura,Y.
Direct Submission
Submitted (02-APR-1999) to the DDBJ/EMBL/GenBank databases.
Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of
Gene Structure 2; 1532-3, Yana, Kisarazu, Chiba 292, Japan
(E-mail:y.nakamura@kazusa.or.jp, Tel:+81-438-52-3935,
Fax:+81-438-52-3934)

```

```

FEATURES
Location/Qualifiers
source
1..84896
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MML2"
/clone_11b="Mitsui P1"
BASE COUNT
27260 a 15039 c 15617 g 26980 t
ORIGIN

```

```

Query Match 13.6% Score 246.8; DB 7; Length 84896;
Best Local Similarity 61.6% Pred. No. 3.1e-45;
Matches 471; Conservative 0; Mismatches 222; Indels 71; Gaps 2;

```

```

QY 707 cagctgaagaggacccagaagtgacacctgtgtcaactgaagaaggtttctctgcgaat 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67675 CAGACCAAAACAGAGAGAAATGCTGCTAGGAAATCTAAGAGATTAAATTCAAAAGAAC 67616

QY 767 tgcagctgcgaacggatactttagtaca-----taattgaagaagtgatttgta 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67615 TTCAATCCGCAACTATGACTTCAGCAGCAAGAAATCTGTCGGAAGAAAGAGGGTTGGAA 67556

QY 821 aggtgtataaaggagacctgtgctgactgtgacactgtgacagtttaaaagccttaagaag 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67555 AAGTGTATTAAGAGTGTGCTTATGATGAGAGTATCAATCGCGGTGAAGATTAAAGATA 67496

QY 881 aacgaacacccaggtggtgagctgcagtttcaacagaagtgaaatgattagcatgctg 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67495 TAAACAAATGCTGTGAGAGGTTCAAGTTTCAGACAGAGCTTGAAAGATGATTAAGCTTGGCG 67436

```

Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63108, USA
e-mail: rwilson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by M. Lohdt, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems such as compressions and repeats; all regions were covered by sequence from more than one subclone

Actual start of this clone is at base position 1 of F17123; actual end is at 134784 of F17123.

```

/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/clone="F17123"
/chromosome="IV"

```

BASE COUNT	43094 a	23410 c	23696 g	44584 t
ORIGIN				

Query Match	11.9%;	Score 216.2;	DB 8;	Length 134784;
Best Local Similarity	59.8%;	Pred. No. 2.2e-38;		
Matches 460;	Conservative 0;	Mismatches 238;	Indels 71;	Gaps

```

Oy      733  cttggtcaactgaagagtttctctgcgagaattgcaagtcgcaagaaacctttagt  792
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 46131  CTGGGAATCTAAGAGCCTTCACATTCAGAGAACTCCATGTTATACAGATGGTTCAGT  46190

```

QY 793 ----accatcttgaaaggtctgtataaagggaagccttgcctgat 846

[illegible]

DB 46251 C C C C A A T C C C C A C C T T C A A C C A T T A A T G G A A C C T C A G G G A T T C A C A G 463

Qy 907 ttccaacagaggtggaatgattagcatgctgtgcatcgaatcttctgcgtctacgt 966

Db 46311 TTTGCTATGGAGCTAGAGATGATTAGCTTAGCTGTTCAATAGAAATCTGCTTCGGTTAATT 463

[illegible]

105

Db 46431 GTCCCTCTAAGCTTAAATGTTAGTATAATGCTTGAACCATCAGTCCAAATCTCATTTT 464

QY 1061 -----aacctcccttgattggcaa 108

D6 46491 CGATGAGTTTAAACCAATTGTGTTTAACTGTGTGAGCTAAACCGGCAATGGACTGGACA 465

[illegible]

Cir 11A) +ccccccca+at+a+ccca+c+c+cca+rt+aaaaaccr+c+ccaaaaata+ta+tt+at+tcaccccacaacaa+t+q 120

Db 46611 ATCCCAAGATCATTCATAGAGATGTAAGGCAGCTAATATCTCTTAGACGAGTGCTTG 466

QY 1202 aggcctgtgtagtgattttgtggttagctagctcatgattacaagatacccatgtta 120

Db 46671 AACGCTGTTGGTGACTTGGACTGCGAAGACCTCCTTAACCAACGCGATTCTCATGTCA 46730
 Qy 1262 caactcgttaagggttaccttggtgacatagctcccgagtaacctcgcgtgaaagt 1321
 Db 46731 CAACCTCGGTCGGTGTAGCTGTGGACACATGACCGTGAATATCTCCACCTGGTCACT 46790
 Qy 1322 catcagaagaagaccgagctgcttggttatgagatatagctcttaagactcattcgtgac 1381
 Db 46791 CTCTCTAGGAAACCGATGCTGTTGGTTCGTACTACTTGTGTCGAGCTCATTAACCGGAC 46850
 Qy 1382 agagagcttgatgctgctgcgcttgagacgagatgagatgattgct 1430
 Db 46851 TGAAGACTCTTGAGTTGGTAAACCGTTAGCCACAAAGGAGACTTACT 46899

RESULT 14
 LOCUS AC002292/2
 DEFINITION Genomic sequence of Arabidopsis BAC F8A5, complete sequence.
 AC002292
 VERSION AC002292.1 GI:2252639
 KEYWORDS HTG.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsis.
 1 (bases 1 to 120787)
 Federpspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
 Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,
 Kim,C., Li,Y., Oj.L.O., Osborne,B., Shen,Y.K., Toriumi,M., Vyotskala,V.,
 Yvostkala,V.S., Yu.G., Ecker,J., Theologis,A. and Davis,R.W.
 Genomic sequence of Arabidopsis
 Unpublished (1997)
 2 (bases 1 to 120787)
 Federpspiel,N.A., Davis,R.W., Conway,A.B., Palm,C.J., Conway,A.R.,
 Kurtz,D.B., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Shin,P.,
 Sun,H., Oj.L.O., Osborne,B., Shen,Y.K., Toriumi,M., Vyotskala,V.,
 Yu.G., Theologis,A. and Ecker,J.
 Direct Submission
 Submitted (05-JUN-1997) Biochemistry, Stanford University/DNA
 Sequencing and Technology Center, 855 California Avenue, Palo Alto,
 CA 94304, USA
 3 (bases 1 to 120787)
 Federpspiel,N.A., Davis,R.W., Conway,A.B., Palm,C.J., Conway,A.R.,
 Kurtz,D.B., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Shin,P.,
 Sun,H., Oj.L.O., Osborne,B., Shen,Y.K., Toriumi,M., Vyotskala,V.,
 Yu.G., Theologis,A. and Ecker,J.
 Direct Submission
 Submitted (05-JUN-1997) Biochemistry, Stanford University/DNA
 Sequencing and Technology Center, 855 California Avenue, Palo Alto,
 CA 94304, USA
 4 (bases 1 to 120787)
 Federpspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
 Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,
 Kim,C., Li,Y., Oj.L.O., Osborne,B., Shen,Y.K., Toriumi,M., Vyotskala,V.,
 Yu.G., Ecker,J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (09-JUL-1997) Biochemistry, Stanford University/DNA
 Sequencing and Technology Center, 855 California Avenue, Palo Alto,
 CA 94304, USA
 5 (bases 1 to 120787)
 Federpspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
 Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,
 Kim,C., Li,Y., Oj.L.O., Osborne,B., Shen,Y.K., Toriumi,M., Vyotskala,V.,
 Yu.G., Ecker,J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (11-JUL-1997) Biochemistry, Stanford University/DNA
 Sequencing and Technology Center, 855 California Avenue, Palo Alto,
 CA 94304, USA
 6 (bases 1 to 120787)
 Federpspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,

gene
 CDS
 complement(8347..9066)
 /gene="F8A5.3"
 /complement(join(8347..8443,8539..8657,8941..9066))
 /gene="F8A5.3"
 /codon_start=1
 /product="putative ADP-ribosylation factor"
 /protein_id="AAB71955.1"
 /db_xref="GI:2462736"
 /translation="MGISFAKLFESRLPAKKEMRIAMGDAACKTITLTKKIGEIDE
 LMDAVLLEAFANKDDLPNPMNAELITDRLGLHSLSKORHWYIQTICATTGGELGDLWL
 SNMINKVSTP"
 complement(9782..14180)
 /gene="F8A5.4"
 /note="Atyps34"
 complement(join(9782..9919,10031..10178,10275..10418,
 10509..10663,10780..10917,10991..11104,11286..11477,
 11813..11936,12047..12182,12300..12395,12475..12622,
 12867..12981,13104..13313,13414..13572,13659..13755,
 13867..13963,14069..14180))
 /gene="F8A5.4"
 /note="location of EST gb|443464"
 /codon_start=1
 /product="phosphatidylinositol 3-kinase"
 /protein_id="AAB71971.1"
 /db_xref="GI:2462752"
 /translation="MGANERFPLSCDINSVPYRIEKLGNLIPVKKSSDGSVSTAE
 EKPRILECALYIDGAFGIDPMRTRLKTGTPPCWMLTILSSKRYDLTAHSQLAT

gene
 CDS
 complement(3171..4444)
 /gene="F8A5.2"
 /complement(join(3171..3389,3471..3605,3706..4029,
 4118..4444))
 /gene="F8A5.2"
 /note="nearly identical to rice water stress induced
 protein gp1D65371537404"
 /codon_start=1
 /protein_id="AAB71970.1"
 /db_xref="GI:2462751"
 /translation="MAPEISVNPMTYSEKARHAPPRAYTFPLAGNDYVKGVGGLAK
 GLRKYSKYPVYAMLPDPEHREILIRSGCYVRELEPPYPPDNQVEEAMAYVYIAT
 SKLRWNFEYSKATVLDADIQVNDIDHFLDSATFAVMDCEKMTWSHSLSI
 GYCOQCEPKVWPEPDESPPPLYLENKGMPSPLETSLSLQTLLETPPYSPFAOD
 FLNMFEEYKVPDLYVNLVLMRMHRENVLEKAVHYCAAGSKPRKPYGEENM
 REDKIMLVDRKMDVYNDESIDEFKSLPADAEETVYKSSITLASVLEPEMTYPAPASA
 "
 complement(8347..9066)
 /gene="F8A5.3"
 /complement(join(8347..8443,8539..8657,8941..9066))
 /gene="F8A5.3"
 /codon_start=1
 /product="putative ADP-ribosylation factor"
 /protein_id="AAB71955.1"
 /db_xref="GI:2462736"
 /translation="MGISFAKLFESRLPAKKEMRIAMGDAACKTITLTKKIGEIDE
 LMDAVLLEAFANKDDLPNPMNAELITDRLGLHSLSKORHWYIQTICATTGGELGDLWL
 SNMINKVSTP"
 complement(9782..14180)
 /gene="F8A5.4"
 /note="Atyps34"
 complement(join(9782..9919,10031..10178,10275..10418,
 10509..10663,10780..10917,10991..11104,11286..11477,
 11813..11936,12047..12182,12300..12395,12475..12622,
 12867..12981,13104..13313,13414..13572,13659..13755,
 13867..13963,14069..14180))
 /gene="F8A5.4"
 /note="location of EST gb|443464"
 /codon_start=1
 /product="phosphatidylinositol 3-kinase"
 /protein_id="AAB71971.1"
 /db_xref="GI:2462752"
 /translation="MGANERFPLSCDINSVPYRIEKLGNLIPVKKSSDGSVSTAE
 EKPRILECALYIDGAFGIDPMRTRLKTGTPPCWMLTILSSKRYDLTAHSQLAT

COMMENT
 FEATURES
 source
 1..120787
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 177..2876
 /gene="F8A5.1"
 join(177..214,632..707,1132..1152,1283..1479,1617..1761,
 1845..2159,2256..2526,2788..2876)
 /gene="F8A5.1"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AAB71954.1"
 /db_xref="GI:2462735"
 /translation="MCCRWVTLRQNDFTGEVSLVFGSDVYVAPITGFQKIYLOI
 LNTLIDLMVKGSTSPGNOQYFAVNAADKPCPTASNLRLKSGLEDYVLRANGLD
 MDCYDFSRHLKXGSGTGOEDKHKRGVGTMTVYVLSLESTQHCSSGNGTEV
 LYRNPFLSPVPHALSKRKIDMKIGLILAVNODDHFLEMDPEPSYQIOIAL
 HWYINQPTROKNSPGISLAKGKIKNALDMRKHSGFLLSSSKRTCTYVDLARS
 AGLTFSSTDDFDGDCSLVSGFQTEVERTVENYIQRKIVYVIGNERKPKQDEAA
 PFLFDESESTFEDEVEDRENSHETDATTWKAYTOFKCLATLMOE"
 complement(3171..4444)
 /gene="F8A5.2"
 /complement(join(3171..3389,3471..3605,3706..4029,
 4118..4444))
 /gene="F8A5.2"
 /note="nearly identical to rice water stress induced
 protein gp1D65371537404"
 /codon_start=1
 /protein_id="AAB71970.1"
 /db_xref="GI:2462751"
 /translation="MAPEISVNPMTYSEKARHAPPRAYTFPLAGNDYVKGVGGLAK
 GLRKYSKYPVYAMLPDPEHREILIRSGCYVRELEPPYPPDNQVEEAMAYVYIAT
 SKLRWNFEYSKATVLDADIQVNDIDHFLDSATFAVMDCEKMTWSHSLSI
 GYCOQCEPKVWPEPDESPPPLYLENKGMPSPLETSLSLQTLLETPPYSPFAOD
 FLNMFEEYKVPDLYVNLVLMRMHRENVLEKAVHYCAAGSKPRKPYGEENM
 REDKIMLVDRKMDVYNDESIDEFKSLPADAEETVYKSSITLASVLEPEMTYPAPASA
 "
 On Jul 11, 1997 this sequence version replaced gi:2251217.
 location/Qualifiers
 1..120787
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 177..2876
 /gene="F8A5.1"
 join(177..214,632..707,1132..1152,1283..1479,1617..1761,
 1845..2159,2256..2526,2788..2876)
 /gene="F8A5.1"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AAB71954.1"
 /db_xref="GI:2462735"
 /translation="MCCRWVTLRQNDFTGEVSLVFGSDVYVAPITGFQKIYLOI
 LNTLIDLMVKGSTSPGNOQYFAVNAADKPCPTASNLRLKSGLEDYVLRANGLD
 MDCYDFSRHLKXGSGTGOEDKHKRGVGTMTVYVLSLESTQHCSSGNGTEV
 LYRNPFLSPVPHALSKRKIDMKIGLILAVNODDHFLEMDPEPSYQIOIAL
 HWYINQPTROKNSPGISLAKGKIKNALDMRKHSGFLLSSSKRTCTYVDLARS
 AGLTFSSTDDFDGDCSLVSGFQTEVERTVENYIQRKIVYVIGNERKPKQDEAA
 PFLFDESESTFEDEVEDRENSHETDATTWKAYTOFKCLATLMOE"
 complement(3171..4444)
 /gene="F8A5.2"
 /complement(join(3171..3389,3471..3605,3706..4029,
 4118..4444))
 /gene="F8A5.2"
 /note="nearly identical to rice water stress induced
 protein gp1D65371537404"
 /codon_start=1
 /protein_id="AAB71970.1"
 /db_xref="GI:2462751"
 /translation="MAPEISVNPMTYSEKARHAPPRAYTFPLAGNDYVKGVGGLAK
 GLRKYSKYPVYAMLPDPEHREILIRSGCYVRELEPPYPPDNQVEEAMAYVYIAT
 SKLRWNFEYSKATVLDADIQVNDIDHFLDSATFAVMDCEKMTWSHSLSI
 GYCOQCEPKVWPEPDESPPPLYLENKGMPSPLETSLSLQTLLETPPYSPFAOD
 FLNMFEEYKVPDLYVNLVLMRMHRENVLEKAVHYCAAGSKPRKPYGEENM
 REDKIMLVDRKMDVYNDESIDEFKSLPADAEETVYKSSITLASVLEPEMTYPAPASA
 "

Query Match

gene
CDS
11.9%; Score 215.4; DB 8; Length 120787;
complement(join(26476..26936,27030..28070,28540..28975))
/gene="F8A5.8"
/gene="F8A5.5"
/gene="F8A5.6"
/gene="F8A5.5"
/gene="F8A5.5"
/note="similar to "Mx" GTP-binding proteins"
/codon_start=1
/protein_id="AAB71966.1"
/db_xref="GI:2462747"
/translation="MYKRVKRVKNSPSIAIAKSKSSKNDVSVYAPAISSYNDRIK
PLIDVDRLRLNVMREGIHLPTIVVGDQSSGKSVLSLAGISLPRQOIGCTVPL
VMRIQSSSEPEIWELENDKVPDEEHIAIRATAVIGSGSDAPLTLVHK
KAGVPLTWDLGITRVVNGOPENIYEOISGMIMEYIEPQESIILNVSATVETT
CESIRMSRVDTGORTLAVTKADMAPGLIOLQTAADPVSIVIGCVARNIGSETY
EARMOBELLFRHPVLSIDEDVIGIPYIAQKMLIQNTKISRLPEITVKINHKETVALEKLPV
TAYELNKLPMVASTGEALMALMDITGSKESLRLIVQGPSETPDDQNNHCTARL
ADMLSQPSLSQAKPKVEAFELMEIKILDECKCVGLPNFIPRSAPLALISLQVHGIQ
DKPVEFINKIWDYIEDVLSVTKARSDNFPQIOSIKRAGRNLISIKIKQSVNRVIE
VEMEKLTDCNPEYMTSWTKTSAQSEFIDAVVKNENIPDYFSVSGFNQVIRK
YHAHLIPADMKRITISWKIVLRIVNLAIVLYOLQSVLNTVEFOKEIVAEWVP
RDGCVKMESESPVYASKREKQNSIKILKESKDAVAIVQNC"
18712..22510
/gene="F8A5.6"
/gene="F8A5.6"
/gene="F8A5.6"
/note="similar to "Mx" GTP-binding proteins"
/codon_start=1
/protein_id="AAB71956.1"
/db_xref="GI:2462737"
/translation="MGSKKHVYRTSSPSIAIVQANPHDNREVPTEAPITISYNDRI
IRPLDVTDLRLNVMREGIOLPTIVVGDQSSGKSVLSLAGISLPRQOIGCTRV
PLVMRIQSSSEPEIWELENDKVPDEEHIAIRATAVIGSGSDAPLTLVHK
NSRNVNGOPENIYEOISGMIMEYIEPQESIILNVSATVETT
THPLSLIDEDVIGIPYIAQKMLIQNTKISRLPEITVKINHKETVALEKLPV
MASTGEALMALMDITGSKESLRLIVQGPSETPDDQNNHCTARLADMLSQVHGIQ
EKPEVTEFLMEIKILDECKCVGLPNFIPRSAPLALISLQVHGIQVKEFEFKITMD
YIEVLSVIAKYSENFPQIOSAIKRAGRNLITIKESHSVERVLEIEMEKLTDCN
PEYMTSWTKADQONITAVILSDGKREHSLTGGGIVKISHLKRYHAHLQOAFD
MKMRIAIVTIVIRIYDLSAIYLOAGGIIOMSESPVASKREKQNSIKILKES
KDAVAIVDQSSGKESLRSKSGISSTTEILTMSDEHIAELISSAARASGSGV
SYPLTLHVRKADVPDLTLVLPETIRVPVNGOTQNHAKNIQDYEHOAT"
23787..25157
/gene="F8A5.7"
/gene="F8A5.7"
/gene="F8A5.7"
/note="similar to "Mx" GTP binding proteins"
/codon_start=1
/protein_id="AAB71957.1"
/db_xref="GI:2462738"
/translation="MGSKKSNDEYIDVAGMSLSIVNPIAPYISSYNDRIPL
DTVRLNVMREGIOLPTIVVGDQSSGKSVLSLAGISLPRQOIGCTRVPLVM
LQSSSEPEIWELENDKVPDEHIAIRATAVIGSGSDAPLTLVHK
VPLDLMVDELPEIRVPVNGOPENIYEOISGMIMEYIEPQESIILNVSATVETT
IRMSROYDKGERTLAVTKADMAPGLIOLQTAADPVSIVIGCVARNIGSETYEEA
RVOEDLFRHPVLSIDEDVIGIIL"
/gene="F8A5.8"
complement(join(26476..26936,27030..28070,28540..28975))

Best Local Similarity 58.2%; Pred.No.3..3e-38;		Matches 479; Conservative 0; Mismatches 251; Indels 93; Gaps 2;	
Qy	701	atgtgccagctggaagaggaccgaagatgcaactgtgtaactgtaagaggtttctctgc	760
Db	109092	ATTTTGCGAACAATATGATCATCAGAACTGATTAGGCGCACTTGAAAGGTATATCATCA	109033
Qy	761	gagaatgcaagctgcaagcgatactt-----tagtacaatacttggaagagtgat	814
Db	109032	AAGACCTTAGATCTGCCCAATCATTTTCAACTCGAAGAACATTTCTCGAAGGCGGAT	108973
Qy	815	ttgtaaggtgtataagggacgctgctgtaagctgtaactgtgaagcgtta	874
Db	108972	ACGGATTTGTGTACAAAGGACACTTAAACGATGGAACCTTGGTGGCTGTCAACGTCCTCA	108913
Qy	875	aagaagaacgcaacacaggtgtgagctgcagttcaacaagaggtgaaatgattagca	934
Db	108912	AGGACTGTACATTCGCGGTGAGAAAGTCCAGTTTCAGACAGAAAGTGAAGTAACTTAAGTT	108853
Qy	935	tggctgtgcatcgaaatctctgcgtctagctggttctctgcatgacacccaagagcgc	994
Db	108852	TGGCTCTTCATCGCAATCTCCTCCGGCTCCGGTTTCTGTAGTACGACACGAGAGAA	108793
Qy	995	ttctgtatatacacaatgctgaagtgctgctggttgaagagcgtcagc	1054
Db	108792	TTTGTAGTACCCCTTACATGCAAAATGGAGAGTGTGCATCAGCTTAAAGGTATGTCAG	108733
Qy	1055	catc-----	1059
Db	108732	TTTCCAAAAAACGTCATATTTCTTCAATGCTTTTTCGACAGTTTAAATGTTCCATCTT	108673
Qy	1059	-----agaacctcccttgattggtgcaactaga	1087
Db	108672	GCTACATTTTGTGTTAGATATATCCGTGGAGAGCCAGCATTTAGACTGTGCGAAGAA	108613
Qy	1088	agaagattgacatagatcgtctaggggtcttcttattgcatgacatggtatccca	1147
Db	108612	AGAAGATAGCGGTTGGGACAGGAGAGCATGTTACCTTACAGAGCAAGGTACCCGA	108553
Qy	1148	gaattatcatcgtgattgattaaagctgcaaatataatttgagcgaagaattgaggtg	1207
Db	108552	AGATTATACCGCAGATGTGAAGACGCTTAACATTTCTTGAATGAGAGACTTCGAAACAG	108493
Qy	1208	tgtgaagttatttgggttagttagctagctcatgattacaagatacccatgttacaactg	1267
Db	108492	TGTTGGTATTTTGGTTAGTAACTTACAGCATTAAGAGACTCTCATCTCACTCAACTG	108433
Qy	1268	ctgtaaaggtacaccttggtgctacatagctcccgatcactctgactggaagtcacatag	1327
Db	108432	CAGTCCGTGAACGTGTTGGCCACATTTGCACTGAGTACTTATCCAGGGTCAGTCTCAG	108373
Qy	1328	agaagaacgattctcttggttagatgattgctttagagctcaacttgcagacagagag	1387
Db	108372	AGAAGCTATGTTGCTTTGGCTTTGGCATTACTTCTCTGAGACTCATTTAGCTCAAGAAAG	108313
Qy	1388	ctttgattctgctgcgcttgcaagcagatgatatgattatglt	1430
Db	108312	CTCTTGATTTTGGAGATCCGCACACAGAAAGGTATGCT	108270
RESULT 15			
LOCUS	AC005170/c	AC005170	110211 bp DNA
DEFINITION	Arabidopsis thaliana chromosome II section 135 of 255 of the complete sequence.		
ACCESSION	AC005170	AE002093	
VERSION	AC005170.2	GI:6598448	
KEYWORDS	HTG.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;			

[illegible]

Db	53985	AAAAAAAAAGTTTCGATACTGCTTTTAGTAAAAAATAATGAATACCTTGAACCTGA	53926
Qy	1055	-----catagaacctccctctgtatctggccaactaggaagagatgtcac	1099
Db	53925	TTTTTGTACTGTTACAGCTAAAGCCAGCGTTGACCTGGAACACAGAGAAAGATATGGCA	53866
Qy	1100	tagagatctgaaggggcttctcttatttgcatagcacatttgcataccaaatattccalc	1159
Db	53865	TTGGAGCTGCAAGAGGGTGTGTTTATCTACACAGAGCAATGGATCCGAATATTACACC	53806
Qy	1160	gtcagtcaaaagctgcacaatatattatgtgacgaagaatttggagctgttgtatgtact	1219
Db	53805	GAGATGTCAAGGACGAAACATTCTCTAGATGAGTATTTTGAAGCAGTTGTGGGGATT	53746
Qy	1220	tttgggttagctaggtcctatgtatgtacaaagatacccatgttacaactcgtcttaagggtta	1279
Db	53745	TTGAGTACGAAAGCTACTACACCCAGAGGATTCACATGTCTACACACCGCGTTAGAGGAA	53686
Qy	1280	ccttgggtctacatagctcccgagtaacctctgactgtgaagaagtcatcagagaagacccgag	1339
Db	53685	CTGTGGTGTCACTTGCACCTGTGATATCTCTCCACCGGTCACTCTCGAATAAACCGATG	53626
Qy	1340	tctttgttatgtgattatgcctcttagagctcattactgcagacagaggctttgactctg	1399
Db	53625	TCTTTGGGTTCCGGTAACTATTGCTAGAGCCTCATCCACAGGAATGAGAGCTCTCGAGTTTG	53566
Qy	1400	ctgcgccttggcaagatgtatgtatgtattt 1430	
Db	53565	GCAAGTCTGTTAGCCAGAAAGGACATAGCT 53535	

Search completed: June 23, 2000, 23:41:42
Job time: 43857 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:48:42 ; Search time 446.21 Seconds
(without alignments)
1017.679 Million cell updates/sec

Title: US-09-180-798-2

Perfect score: 1815
Sequence: 1 gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N.Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1803	99.3	1814	1	Daucus carota SERK
2	933.2	51.4	2089	1	Arabidopsis thaliana
3	560.2	30.9	6695	1	Daucus carota SERK
4	270.8	14.9	4081	1	Arabidopsis thaliana
5	143.4	7.9	3842	1	Tomato Xa21 clone
6	142.4	7.8	4104	1	Arabidopsis sterol
7	142	7.8	3045	1	Maise Xa21 gene DT
8	140	7.7	1554	1	Tomato RRK gene cl
9	140	7.7	3293	1	Tomato Xa21 clone
10	138.4	7.6	3176	1	Arabidopsis thaliana
11	121.2	6.7	2640	1	S-receptor
12	115.4	6.4	2833	1	S-receptor kinase
13	114.8	6.3	2685	1	SRK-A10 cDNA. DNA
14	114.8	6.3	2685	1	SRK-A10 gene mutan
15	112.6	6.2	9295	1	Arabidopsis thaliana
16	110.4	6.1	2749	1	S-Locus receptor (
17	107.2	5.9	2780	1	Brassica sp. S-rec
18	103.4	5.7	788	1	Arabidopsis thaliana
19	103.4	5.7	1063	1	Arabidopsis thaliana
20	101.8	5.6	981	1	Arabidopsis thaliana
21	99.6	5.5	2443	1	Tomato Pto protein
22	99.4	5.5	1269	1	ZmPK1 homologue pr
23	95.4	5.3	894	1	Arabidopsis thaliana
24	92	5.1	2704	1	Class II S-receptor
25	88.8	4.9	2749	1	S-receptor kinase
26	87.6	4.8	5733	1	Arabidopsis thaliana
27	84.2	4.6	1106	1	Arabidopsis thaliana
28	76	4.2	960	1	Arabidopsis thaliana
29	73	4.0	3573	1	Protein kinase gen
30	73	4.0	6471	1	Partial tomato pat
31	70.2	3.9	3979	1	Tomato pathogen re
32	70.2	3.9	3979	1	CF-5 pathogen resi
33	70.2	3.9	4123	1	CF-5 pathogen resi
34	64.6	3.6	3541	1	CF-5 pathogen resi

35	62.6	3.4	19639	1	X23524	O. longistaminata
36	59.6	3.3	7204	1	X23527	O. longistaminata x
37	58.6	3.2	277	1	V60314	Resistance marker
38	58.6	3.2	3921	1	T31300	Rice Xa21 disease
39	58	3.2	13340	1	X23522	O. longistaminata
40	57.2	3.2	8416	1	X23523	O. longistaminata
41	55.6	3.1	5940	1	X23526	O. longistaminata x
42	54.8	3.0	2075	1	T49435	Tomato polygalactu
43	54.2	3.0	960	1	T35826	Protein kinase gen
44	54	3.0	9424	1	X23525	O. sativa Xa21 gen
45	53.8	3.0	6256	1	T31299	Rice Xa21 disease

ALIGNMENTS

RESULT	ID	Sequence	Score	Length	DB	Match	Query Match	Length	ID	Description
1	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Daucus carota SERK
2	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
3	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Daucus carota SERK
4	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
5	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Daucus carota SERK
6	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
7	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Tomato Xa21 clone
8	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis sterol
9	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Maise Xa21 gene DT
10	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Tomato RRK gene cl
11	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Tomato Xa21 clone
12	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
13	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	S-receptor
14	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	S-receptor kinase
15	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	SRK-A10 cDNA. DNA
16	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	SRK-A10 gene mutan
17	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
18	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	S-Locus receptor (
19	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Brassica sp. S-rec
20	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
21	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
22	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Tomato Pto protein
23	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	ZmPK1 homologue pr
24	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
25	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Class II S-receptor
26	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	S-receptor kinase
27	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
28	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Arabidopsis thaliana
29	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Protein kinase gen
30	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Partial tomato pat
31	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	Tomato pathogen re
32	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	CF-5 pathogen resi
33	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	CF-5 pathogen resi
34	V06571	gacaataaccattgaatat.....ttaaaaaaaaaaaaaaaaaa	99.3%	1814	1	1814	99.3%	1814	1	CF-5 pathogen resi

QY 121 aattacatgacgttcaatgagcttacccttgacaataatgaggttcttaataatgag 180
 |||||
 Db 121 AATTACATGCACTGATGCTTACCTTACCAATATGGGGTCTTATGACATTTGAG 180
 QY 181 cttaacagaataaataagtaggaccacattccctagtagtctctgggaatctgacaattg 240
 |||||
 Db 181 CTTTACAGCAATATACATATAGTGGACCAATCTCTAGTGTCTTGGGAATCTGACAAATTTG 240
 QY 241 gtagagcttggaccctataatgaaatgactctctggaccctataccggacacattagaaag 300
 |||||
 Db 241 GTGAGCTTGACCTTATACATGATAGCTTCTGTGACCTATACGGACACATTTAGGAAG 300
 QY 301 cttaacagcttaagattcttgctgctcaacaaacacccctctgtgccaattcccaatg 360
 |||||
 Db 301 CTTACAAAGCTTAAAGATTTCTGCTCAACAAACACAGCCCTCTGCTGCTCAATTCATG 360
 QY 361 tcaactgacaataatatacaactctcaagtcctggaattcacaacaaatcggctatcagga 420
 |||||
 Db 361 TCACGTACATATATACAACTTCAAGTCCCGATTTATCAAAATCGGCTATCAGGA 420
 QY 421 ccaagtaacggataatgagctcaattctctgtttacacatacagtttcccaataattg 480
 |||||
 Db 421 CCAAGTACCGGATATGGCTCATTTCTTGTGTACACCTATCAGTTTGGCCAAATATTTG 480
 QY 481 aattatgtggaccccttaactgagagccctgcccctggaatcccccattctcgccaaca 540
 |||||
 Db 481 AATTATGTGGACCCCTAATCTGGAGGCCCTCCCTGATCTCCCCCATTTTCCGCACCA 540
 QY 541 cctccgttcaatcccaatcaacagtaacagcctccaggaacaaatgtgccaatgagct 600
 |||||
 Db 541 CCTCCCTTATCCCAACATCAACAGTACAGCTCCAGGACAAATATGGTCCCACTGGAGCT 600
 QY 601 attgagcggagagtagctgctgctgctgcttactgagtttgcgcgcgcgcgaatgacatt 660
 |||||
 Db 601 ATTGCTGGGGAGATGCTCTGTGCTCTTACTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 661 gcatgtgtgaggaagaaacacgaggaacattcttctgtagtgcagctggaagagac 720
 |||||
 Db 661 GCATGTGTGGGGAGGAAGAAACCGGAGAACATTTCTTGTGATGTGCCAGCTGGAAGAGAC 720
 QY 721 ccaagaagtgacacttgtaactgaaagaggtttctctgaggaattgcaagtcgaag 780
 |||||
 Db 721 CCAGAAGTGCACCTTGTGTAACCTGAGAGAGGTTTCTCTGCGAATTTCAAGTCCCAAG 780
 QY 781 gatactttgatacacttactgaaagaggtgagattgtgaaagtgtataaaggaagcctt 840
 |||||
 Db 781 GATACCTTTTATGATCACTTACTTGGAGAGAGTGGATTTGTAAGTATTAAGAGGAGCCTT 840
 QY 841 gctgagtgtcactgtgtagcaagttaaaggcttaagaagaagaacacacaggtgtag 900
 |||||
 Db 841 GCTGATGGCTCACTTGTAGCAGATTAAAGCTTAAAGAAAGAAACGACACAGGTGTGAG 900
 QY 901 ctgcaatttcaacacagaggtggaatgattgacatgctgctgcatcgaaatctctgct 960
 |||||
 Db 901 CTGCAATTTCAAACAGAGGTGGAATGATTAGCATGGCTGTCTCATCGAATTTCTGCTG 960
 QY 961 ctacgtgtttctgcatgacacacacacagacgagcttctgtatatacacaatgagctat 1020
 |||||
 Db 961 CTACGTGTGTTTCTGATGACACACACACAGAGCGGCTTCTGTATATCAATACATGCTAAT 1020
 QY 1021 ggaagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1080
 |||||
 Db 1021 GGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 QY 1081 actaggaagagattgcaactagatctgtagagggcttcttcttattctgacatgct 1140
 |||||
 Db 1081 ACTAGGAAGAGATTGCACTAGATCTGTGAGGGGCTTCTTATTTGATGATACATTTGT 1140
 QY 1141 gatcccaagattatcatcgt 1200
 |||||
 Db 1141 GATCCCAAGATTATCCATCTGT 1200
 QY 1201 gagcgctgttaggtgatttgggttagcagctcatgattacaagaatacccatgct 1260
 |||||

Db 1201 GAGGCTGTTGATGATGATTTGGGTTAGCTAGCTCATGATTAACAAGATACCAATGTT 1260
 QY 1261 acaactgtctgaaggggtgactcttggtgacataagctcccgagtaacctcgcagtgaag 1320
 |||||
 Db 1261 ACAACTGCTGTAAAGGGGTACTTGGGCTTACTAGCTCCAGATACCTCTGATGAGAAAG 1320
 QY 1321 tcatcagaagaacacagatgcttctgtgttagtggttctgcttagtggtgcctatactgga 1380
 |||||
 Db 1321 TCATCAGAAAGACCGAGTGTGTTGGTTAGGATATGCTCTTACGCTCATATCTGGA 1380
 QY 1381 cagagagctttgactctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1440
 |||||
 Db 1381 CAGAGAGCTTTGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 1441 gttaaagcctttgaaagaaagaaagtgtgagatgctgtgctgactgacctgagagac 1500
 |||||
 Db 1441 GTTAAAGCCTTTTAAAGAAAGAAAGTGGAGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1501 aattacattgacagaagattgagagcttataaagtagaattactctgacacaggt 1560
 |||||
 Db 1501 AATTACATTGACACAGAGTTGAGCAGCTTATTCAGATGATATCTGTACCCAGGCT 1560
 QY 1561 tcgccaatgagacggtcctaagatgtcagaggttagtccgaatgcttgaagtgatgagctt 1620
 |||||
 Db 1561 TCGCCATGAGACGGCTTAAAGATGTCAAGATGTCCGAATGCTTGAAGGTGATGCGCTT 1620
 QY 1621 gcaagaagaatgagagagatgagcaaaaagtagaattcatcaataagaacgttagaattgct 1680
 |||||
 Db 1621 GCAGAAAGTGGAGACAGTGGGCAAAAAGTAAAGTCAATCATCAAGACGTGAATTAAGCT 1680
 QY 1681 ccacatcgacttctgaaatgatacctagactcgacagataactgcatgcttctgaaat 1740
 |||||
 Db 1681 CCACATCGAATCTTGAATGATCTCTAGATCGACAGATTAATCTGATGCTTTGAATTA 1740
 QY 1741 tctgtgtcaagatgaacagcatataaagtgtgaatgaatgaatgaatgaatgaatgaatgaat 1800
 |||||
 Db 1741 TCTGTGTCAAGAT -AACAGCATATTAATGTGAATGAATTAATTTTATGTTTAA 1799
 QY 1801 aaaaaaaaaaaaaa 1815
 |||||
 Db 1800 AAAAAAAAAAAAAA 1814
 |||||

RESULT 2
 V06591
 ID V06591 standard; cDNA to mRNA; 2089 BP.
 AC 03-AUG-1998 (first entry)
 DE Arabidopsis thaliana SERK gene.
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
 OS Arabidopsis thaliana.
 FH Key location/Qualifiers
 FT CDS 195..2072
 FT /tag= a
 FT /product= SERK protein
 PN W09743427-A1.
 PD 20-NOV-1997.
 PE 13-MAY-1997; E02443.
 PR 14-MAY-1996; GB-010044.
 PA (NOV) NOVARTIS AG.
 PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;
 DR WPI; 98-086529/08.
 PT Production of apomictic seeds - useful in plant breeding
 PS Claim 27; Pages 91-95; 123pp; English.
 CC The sequence is that encoding SERK, a putative receptor kinase.
 CC It may be used as part of a method of producing apomictic seeds.
 CC comprising: (a) transforming plant material with a nucleotide
 CC sequence encoding a protein which in active form in a cell or
 CC cell membrane renders the cell embryogenic; (b) regenerating
 CC the transformed material into plants or carpel-containing

CC plant parts: and (c) expressing the sequence in the vicinity
CC of the embryo sac. The apomictic seeds and embryos thus produced
CC can be developed into plant progeny. This is useful in plant
CC breeding programs. Controllable and reproducible apomixis provides
CC many advantages in plant improvement and cultivar development in
CC the case that sexual plants are available as crosses with the
CC apomictic plant. Apomixis provides for true-breeding, seed
CC propagated hybrids and could shorten and simplify the breeding
CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;

Query Match 51.4%; Score 933.2; DB 1; Length 2089;
Best Local Similarity 74.4%; Pred. No. 1.6e-229;
Matches 1191; Conservative 0; Mismatches 403; Indels 6; Gaps 1;

QY 175 ttgagcttacagcaatacaataagtgagcaattccctagatcttggaattgaca 234
DB 486 TTGAGCTTACAGTACCAACATCACTAGCGCCCATTCCTAGTAACTTGGAAATCGACA 545
QY 235 aatttgagcttgagcctatacatgaatagctctctggaactataccgagacatia 294
DB 546 AACTTAGTGAAGTTGGATCTTACTTAACAGCTTCGCGCTCTTCGGAATCATTTG 605
QY 295 ggaagcttcaagagctaaatctctgctctcaacaacaagcctctctgccaatt 354
DB 606 GGAAGCTTCAAGCTGAAGATTCCTCGGCTTACACAAACAGTCTCACTGGGTAAAT 665
QY 355 ccaatgactgactgaataatacaactctcaagtcctgactgattatcaacaatcgcta 414
DB 666 CCAATGCACTGACCAATATTAATACCTTCAAGTGTAGATATCAATAATACAAACATC 725
QY 415 tcaagacagtlaccggaataagtcactatcttctgtttacactatcagtttgcaat 474
DB 726 TCTGTTCACTTCCTGACATGGCTCCTCTCACTCTTACACCCATCAGTTTGGTAAAT 785
QY 475 aatttgattatgtagaccgtaactgaagcgcctgacctgatactcccatcttg 534
DB 786 AACTTAGACCTATGTGACCTGTTACAAAGTCAACCCATGCTCGCATCTCCCGTTTCT 845
QY 535 ccaccacccgttcatccaccatcaacagtlacagcctccaggaacaaatgtgccat 594
DB 846 CCTCCACACACTTTTATTAACAACCTCCGCCAGTTTCCACCCGAGTGGGTATTAAC 905
QY 595 ggaactatgctggggagtagctgctgctgcttactgctgctgacactgcaatg 654
DB 906 GGAGCAATAGCTGGGTGAGTGGCTGACAGTGTGCTTGGCCCTTGGTGTCTCGCAAT 965
QY 655 gcaattgcatggtggggagagaacacggcgagaaacattcttcatgctgcaagctga 714
DB 966 GCCCTTGTCTGGTGGGACACGAAGAGCCCACTAGATATTTCTTCGATGTCCTGCGGA 1025
QY 715 gagagaccagaagtgcaccttggtcaactgaagaggtttctctgcaagattgcaagtc 774
DB 1026 GAAGATTCACAAGTCAATCTGGGACACTCAAGAGTTCCTTGGCGGAGTCAACAAG 1085
QY 775 gcaacgatactttagt-----accatacttggaagaggttgattggtlaagtgat 828
DB 1086 GCGAGTATGGGTGTAGTAAACAAGAACATTTGGGACAGAGTGGGTGTGGAAAGTCTAC 1145
QY 829 aagggaagccttgctgtagctcaactgtgacgtttaaaggctttaaagaagaacaa 888
DB 1146 AAGGAGAGCTTGGGACGGAACCTTGTGTCTCAAGAGACTGGAAGGAAGACGAAC 1205
QY 889 ccaaggtgtagagctgcagtttcaacagaggttggaatgattagcagtgctgcacgaa 948
DB 1206 CCAAGTGGAGAGCTCCAGTTTCAACAAGAAAGTAGATGATTAAGTATGAGAGTTCA 1265
QY 949 aatctctgctctacgltgltctgcatgacacaaacagagcggtcttctgatatca 1008

DB 1266 AACCTGTGAATATACAGAGGTTCTGTATGACACCGACGAGAGATGCTTGTATCCT 1325
QY 1009 taactggtataggaagtgtgtgctgttttaagaagagtgtaacgaataaacctccc 1068
DB 1326 TACATGGCAATGGAAGTGTGCTTGTCTGACAGAGAGCCACCGTACAAACCTCCG 1385
QY 1069 ctgattgccaactagaagaagattgcaactagatcctcctaagggtcttctatttg 1128
DB 1386 CTGTATTGGCCAAAGCGGAAAGAAATGCGGTAGCGTACACTGAGGTGTGTCTACCTA 1445
QY 1129 catgaacatgtgatcccaagattatcatcgtgtagtgaagaagctgaatatattg 1188
DB 1446 CATGATCACTCGATCCGAAGATCATTCACCGTACGCTAAAGACGAACATCTCTTA 1505
QY 1189 gacgaagattgagcctgtgtgtagtattggtgttagctagcctcagagattacaag 1248
DB 1506 GACGAAGAAATTCGAGCGGTGTGTGAGATTTCCGGTTGGCAAGCTTATGACTATAA 1565
QY 1249 gataccatgttaacactgctgttaagggtatccttggtctacatagctcccgatctc 1308
DB 1566 GACACTCAGTACAGACAGAGTCCGTGGCACATCGGTACATCGCTCCAGATATCTC 1625
QY 1309 tgcactggaagtcatacagaagaccgagtctcttggtatggaattatgctcttagag 1368
DB 1626 TCAACCGGAAATCTTCAGAGAAACCGAGCTTTCGATACGAATCATGCTCTAGAA 1685
QY 1369 ctcaattactgacagagagcctttagatctgtctgctcttgcaacgtagtattg 1428
DB 1686 CTAATCAGAGACAAAGAGCTTTCGATCTCGCTGGGTAGCTAACGACGAGCTGTATG 1745
QY 1429 ttgttgtagctgtttaaagcctttagaagagaanaagttgagatgctgctgctatcc 1488
DB 1746 TTACTTGACTGGGTGAAGATGTGTGAAGAGAAAGAAAGCTAGAGATGTAGTGAATCA 1805
QY 1489 gaactggaacaaattacattgacacagaagltgagaagcttattcaagtagaattac 1548
DB 1806 GATCTTCAAAACAACTACGAGAGAGAGAACTGGAAACAGATGATACAGTGGCTTCT 1865
QY 1549 tgtaccagaggttgccaatgagcagcgtcgaagtgtgcaagagtgccgaatgctgaa 1608
DB 1866 TGCACGCAAGGATCCCAATGGAAGACCAAAAGATGTCTGAAGTGTGAAGATCTGGA 1925
QY 1609 ggtgtagccttgcaagaagttggaacagtggtgcaaaaagttagaagtcacatcaag 1668
DB 1926 GGAAGTGGGCTTGGCGGAATAATGGACCAATGCCAAAAGTGTGATTTGAGGGAAG 1985
QY 1669 gtgaattagctccacatcgaaactctgnaatgtagtccctagaactcgacagataact 1728
DB 1986 ATTGATTGAGTCTTAATCTTAACCTGATGATGATGATCTTACTTACATTAATG 2045
QY 1729 gctttgattatctcgtgccaaagataacagcatataaa 1768
DB 2046 GCCGTTGAGTTATCTGTCCAAAGTAAAAAATAAAAAA 2085

RESULT 3
V06570
ID V06570 standard; DNA; 6695 BP.
AC V06570;
DR 03-AUG-1998 (first entry)
DE Daucus carota SERK gene.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; ds.
OS Daucus carota.
FH Key
FT Location/Qualifiers
FT CDS 3696..6620
FT /*tag= a
FT /*note= "contains introns"
FT 3731..3802
FT /*tag= b
FT /*number= 1
FT Intron 3851..3979


```

FT      /*tag= f
FT      /number= 5
FT      4642. .4757
FT      /tag= 9
FT      /number= 6
FT      4890. .4967
FT      /tag= h
FT      /number= 7
FT      5295. .5803
FT      /tag= 1
FT      /number= 8
FT      6197. .6339
FT      /tag= j
FT      /number= 9
PN      WO9743427-A1.
PD      20-NOV-1997.
PF      13-MAY-1997; E02443.
PR      14-MAY-1996; GB-010044.
PA      (NOVS ) NOVARTIS AG.
PI      De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
PI      WPI: 98-086529/08.
DR      P-PSDB; W47017.
PT      Production of apomictic seeds - useful in plant breeding
PS      Claim 26; Pages 64-67; 123pp; English.
CC      The sequence is that encoding SERK, a putative receptor kinase.
CC      It may be used as part of a method of producing apomictic seeds
CC      comprising: (a) transforming plant material with a nucleotide
CC      sequence encoding a protein which in active form in a cell or
CC      cell membrane renders the cell embryogenic; (b) regenerating
CC      the transformed material into plants or carpel-containing
CC      plant parts; and (c) expressing the sequence in the vicinity
CC      of the embryo sac. The apomictic seeds and embryos thus produced
CC      can be developed into plant progeny. This is useful in plant
CC      breeding programs: controllable and reproducible apomixis provides
CC      many advantages in plant improvement and cultivar development in
CC      the case that sexual plants are available as crosses with the
CC      apomictic plant. Apomixis provides for true-breeding, seed
CC      propagated hybrids and could shorten and simplify the breeding
CC      process so that selfing and progeny testing to produce and/or
CC      stabilise a desirable gene combination could be eliminated.
CC      Apomixis allows plant breeders to develop cultivars with
CC      specific stable traits for such characteristics as height,
CC      seed and forage quality and maturity.
SQ      Sequence 4081 BP; 1120 A; 770 C; 785 G; 1406 T;

Query Match      14.9%; Score 270.8; DB 1; Length 4081;
Best Local Similarity 64.6%; Pred. No. 7.9e-60;
Matches 504; Conservative 0; Mismatches 157; Indels 119; Gaps 2;

QY      707 cagctgaagaagagaccagaagatgcacctgtgcaactggaagaggtttctctgcgagaat 766
DB      3302 CAGCCGAGAAAGATCCAGAGATTCATCTGGACAGCTCAAGAGGTTTCTTTCGGGGAGC 3361
QY      767 tgcgaatcgcaaggaactttagt-----accatacttggaaagaggtgattgtgta 820
DB      3362 TAAAGATGGGAGTATGGGTGTAGTAACAAGAACATTTTGGGCGAGGTGGGTTTGGGA 3421
QY      821 aggtgtataagagagccttgcgtgactgctcactgtgacattgacaataaaggtcctaagaag 880
DB      3422 AAGTCTAACAGAGGAGCGCTTGGCAGAGCAACTTGTCTCTCAAGAGACTGGAAGAG 3481
QY      881 aacgaacacagagtgctgagctgcagttccaacagaggtggaatgattagcatgctgt 940
DB      3482 AGGAACTCCAGCTGAGAGCTCCAGTTTCAACACAACTAAGATGATGATATGATGAGAG 3541
QY      941 tgcatacgaatctcttcgcgtctacgtgctgttctgcacgaccacaagagcggtctcttg 1000
DB      3542 TTTCATGCAAACTGTGTAAGATTACGAGGTTTCTGTTATGACACCGACCGAGAAATTCCTTG 3601
QY      1001 tatatcatcatcgtgctgaatgaaagtgttgcgtcgtgtttaag-----1044
DB      3602 TGATCTCTACATGCGCAATGGAAGTGTGCTCTGCTCAGAGGTAAACAACTAAACA 3661

```

```

QY      1044 -----1044
DB      3662 TTAAACATCTGTGCTCTCTCAATTACTTTGACGTGAAGTGTTTTTCATGTTTCC 3721
QY      1044 -----agagcgtacgcatcagacctcc 1067
DB      3722 TTATGGGTTTCATATTGTGGTTACACTAATGACACAGAGAGCGCACCGTCACACCTCC 3781
QY      1068 ccttgattggccaactaggaagagatgacatgacatctgctagggggtcttctattt 1127
DB      3782 GCTTGATTGGCCACGCGGAAGAGAAATCGCGCTAGCGCTCAGCTGAGTTGTCTTACCT 3841
QY      1128 gcatgacatgtgtatcccaagattatccatcgtgataagctgcaatatattatt 1187
DB      3842 ACATGATCTACGTGATCCGAAGATCATTCACCGTAGCTAAACACACAACTCCCTT 3901
QY      1188 ggaagaaagattgagagctgtgtgtagtatttgggttagctagctcattgattcaaa 1247
DB      3902 AGACGAGAAATTCGAAGGGGTTGTGAGATTTGCGGGTTGGCAAAAGCTAATGAGACTTAA 3961
QY      1248 gatacccatgttacaactgctgtgaaggggtacattgggtacataagctccgagtaact 1307
DB      3962 AGACACTCAGTGACAAACGACAGTCCGTGGCACCATCGGTCCATCCCTCAGATATCT 4021
QY      1308 ctgcagctggaagatcatcagaagaagccgatgctcttggattatggattatgctttaga 1367
DB      4022 CTCACCGGAAATCTTCAGAGAAACCGAGCTTTTCGATACGGAATCATGCTTTAGA 4081

RESULT 5
X23533
ID      X23533 standard; DNA; 3842 BP.
AC      X23533:
DE      17-JUN-1999 (first entry)
KW      Tomato Xa21 clone TRK2 DNA fragment.
KW      Xa21; receptor kinase-like protein; multigene family; RRR; rice; TRK2;
KW      plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
OS      Lycopersicon esculentum.
PN      MO9090151-A2.
PD      25-FEB-1999.
PF      17-JUL-1998; U14841.
PR      13-AUG-1997; US-910386.
PA      (REGC ) UNIV CALIFORNIA.
PI      Hubert SH, Richter T, Ronald PC, Song W, Szabo V,
PI      Wang G;
DR      WPI: 99-204431/17.
DR      P-PSDB; W23533.
PT      New RRR polynucleotides and nucleic acid constructs - used for
PT      generating transgenic plants resistant to Xanthomonas
PS      Claim 14; Page 61-62; 67pp; English.
CC      This invention describes a method for conferring disease resistance in
CC      plants. The invention describes the use of novel genes and proteins
CC      belonging to the Oryza longistaminata and Oryza sativa receptor
CC      kinase-like protein (RRK) Xa21 multigene family. Such genes from
CC      cassava, maize and tomato are also described. The genes and proteins can
CC      be used for enhancing resistance to Xanthomonas in a plant, preferably in
CC      rice or tomato.
SQ      Sequence 3842 BP; 984 A; 719 C; 894 G; 1244 T;

Query Match      7.9%; Score 143.4; DB 1; Length 3842;
Best Local Similarity 53.4%; Pred. No. 3.2e-27;
Matches 350; Conservative 0; Mismatches 296; Indels 9; Gaps 2;

QY      800 ttgaaagagtgattgtgtaaggtgtataagagccttgcgtgactgactgtag 859
DB      2712 TAGGAGATGAGGTTTGGAGCAACATTAAGACGGAATTTCACACGAGTCTCTAGG 2771
QY      860 cagttaaagcgtttaaagaagaagacacacagcaggtgtgtgagctgagtttcaaacagag 919
DB      2772 CAGTAAAGCGACTTGTGTTAGAGACGTTTTCAGGGGATTT---CAACAGTTGATGACAGAA 2828

```

Qy	920	tggaatagattagatgatgacgctgctgcatccgaattcttcgcttctaogtggcttctgcaga	979
Db	2829	TCAGACCTCTGGGGAGGCTTCGACATCCAAACCTCGTAACCTGTAGTAGGATATCATATA	2888
Qy	980	caccacacagagcgagctctctctgtatcatcacatgagctaatggaagtgttcgctgltt	1039
Db	2889	GTGAACACGAATAGTTCTGATCTATTAACATATTGCCAGGTGGTATTTGGAAGTTTA	2948
Qy	1040	taagaagcgctcagccatcagaaacctcccttgatttgcccaactggaagagatggac	1099
Db	2949	TTCAAGGAGAGCTCTACAAAGG-----CTGTGACCTGGAGGGTTCTTCCAAAGATTGTT	3002
Qy	1100	taggactcgtcaaggggcttctcttatgtcatcaccattggtatcccaagattatcac	1159
Db	3003	TGGATGTAGCCCGCGACACTGCTTACTCGATGATTAAGTGTACACAGCTGTGCTTATC	3062
Qy	1150	gtgatgtaaaagctgcacaatatatttggacgaagaatttgaagcctgtttagtgat	1219
Db	3063	GTGATGTGAAGCCACGACGACMACTTTATTGTGATGAGAGATATATGATTTATTTGATT	3122
Qy	1220	ttgggcttagcctgaagctcatgattacacagatcacccatgttcaactgcttgaagggt	1279
Db	3123	TGTGTTTGGCTAAGTAACTGTGGAACTTCAGAGACCATGCACTACTGTTGGCGGGAA	3182
Qy	1280	ccctgggcatcatagctcccgagtaacctgcagctggaagaatcatcagaagaacagtg	1339
Db	3183	CTTTGGATATGTCCTCCGAAATAGCCATGTGCTCCGCGCTCCGGAACAAGGCTGATG	3242
Qy	1340	tctttggttaatgagatattcctcttagagctaatcactgacagagagattttagctg	1399
Db	3243	TCTACAGTATAGGGGTTGTGTGCTTGAGTTAATATCAATAAGAAAGCACTAAGTCCGT	3302
Qy	1400	ctcgacctgcagacgatgatgatgtatgttcttgagattgggttaaaagcccttt	1454
Db	3303	CTTCTCTCTTAAAGAAATGATTCATATTTGTACTTGGGATGCAAGCTCTTT	3357

RESULT	6
ID	X07356 standard; DNA; 4104 BP.
AC	X07356;
DT	21-MAY-1999 (first entry)
DE	Arabidopsis steroid receptor Bin1 DNA.
KW	Bin1; steroid receptor; receptor kinase; transgenic plant;
KW	brassinosteroid; disease resistance; crop protection;
KW	contraceptive; ss.
OS	Arabidopsis thaliana.
FA	Location/Qualifiers
FT	CDS
FT	97. .3667
FT	/*tag= a
FN	W09859039-A1.
PD	30-DEC-1998.
PR	24-JUN-1998; U13100.
PR	24-JUN-1997; US-881706.
PA	(SALK) SALK INST BIOLOGICAL STUDIES.
PI	Chory J, Li J;
DR	WPI: 99-081275/07.
DR	P-PSDB; W97819.
FT	New receptor kinase Bin1 involved in brassinolide signalling -
PT	useful for promoting increased yield and disease resistance in
PT	plants and for modulating oocyte maturation
PS	Claim 8: Page 49-52; 72pp: English.
CC	This DNA sequence codes for a novel plant steroid receptor kinase,
CC	designated Bin1 (see W97819), which is involved in the pathway for
CC	the synthesis of the plant steroid hormone, brassinolide. 18 New
CC	Arabidopsis dwarf mutants were identified that lacked the ability
CC	to respond to brassinolide, and were named bin mutants. The bin1
CC	mutations were used to map the gene to a small interval on
CC	Arabidopsis chromosome 4. Bin1 was cloned by map-based cloning.
CC	The bin1 polynucleotide was identified within this interval by
CC	sequencing the wild-type and mutant alleles of this nucleic acid.
CC	Overexpression of Bin1 in transgenic plants provides plants
CC	characterised as having enhanced disease resistance, increased

CC Plant yield or vegetative biomass and increased seed yield.
CC Expression of Bn1 may also increase resistance to pestilicides.
CC Inhibition of Bn1, e.g. using antisense oligonucleotides (AON), is
CC used to render plants male-sterile, and to reduce their stature or
CC yield, e.g. for creating dwarf varieties. Since Bn1 homologues
CC may be involved in regulation of the menstrual cycle and uterine
CC function, Bn1, antibodies and AON may be useful as contraceptives,
CC for improving success of in vitro fertilisation and to prevent
CC premature labour. Transgenic animals are also provided, and are
CC models for studying steroid-receptor interactions or can be used
CC to screen for therapeutic agents. 870 C; 930 G; 1243 T;
SQ Sequence 4104 BP; 1061 A; 870 C; 930 G; 1243 T;

Query Match	7.8%;	Score 142.4;	DB 1;	Length 4104;
Best Local Similarity	53.6%;	Pred. No. 5.9e-27;		
Matches 343;	Conservative 0;	Mismatches 291;	Indels 6;	Gaps 2

OY	800	ttggaagaagtgatgttggtcaaaagtgtataaggaagccttgcgtatgctctatgag	859
Db	2762	TTTGCTTGGGTGGGTTTGGAGATGTTTACAAACCGATTTTGAAAAGTGGAAAGCGCGGTGG	2822
OY	860	cagttaaaagcgtttaaagaagaacgaacacacagctgtgtagctgcagtttcaacaagag	919
Db	2822	CTATCAAGAAACTGATTTCAATGTAGGGGTCAA---GGTATATAGAGTTCAATGGGGAGA	2878
OY	920	tggaaatgatatagcatacgtgtgtcatcgaacatcttcctgcgtcaegtttcctcatga	979
Db	2879	TGGAACCATGGGGAACATCAAAACATCGAATCTTGTGCTCTTCTGTGTTTTCGCAAG	2938
OY	980	caccaacagagcggcttcctgtatataccatacctgacgtcctaagtgtgtcgtgtgt	1033
Db	2939	TTGGAGACGACCGGCTCTTCTGTTATATAGAGTTATGAAATATGCAAGTTAGAAATGTTT	2999
OY	1040	taagagagcgtcagccatcagaaacctcccttgatttgcgcaactaggaagagattgcac	1099
Db	2999	TGCAAGAACCCCAAGAAAGTGGGGTGAAACTTAAATTGTCCACACGCGGAGAGATTGGA	3058
OY	1100	tagagatctgtcagaagggcttctctatttgcatagcattgtatcccaagatataccatc	1155
Db	3059	TAGGATCAGCTAGAGGGGCTTGTCTTCCCTTACCACAACTGCAGTCCGCTATATATCCACA	3118
OY	1160	gtgatgtaaagcgtcgaatatatattatggacgaagaatttagagcgtgtgttagttgat	1218
Db	3119	GAGCATCAAAATCCAGTAATATGTGTGCTTATATAGAAATTTGGAAAGTCCGGTTTCAGATT	3178
OY	1220	ttaggttagtaggtcctcatatgcatatacaagaatacccaatgta--caactgtctgaaggy	1276
Db	3179	TTGGCATGGCGAGGCTGATGATGCCATGATACGCATTTAAACGCTACATATTAGCTG	3233
OY	1277	gtaccttgggtataatgatgctcccgagacctctgcagcttggaaagtcatagaagaagcgy	1333
Db	3239	GTACACCGGGTTTACGTTCTCTCCAGATATTACCAAAAGTTTCAGGTCTTCAACAAAAGAG	3298
OY	1337	atgtcttgggttcaaggagattatgctcttaagctcattacttgcagcagagagcctttgag	1396
Db	3299	ACGTTTATAGTTACGATGTGTCTTACTATCGAGCTACTCAAGGGTTAAAGGCCAACAGGATT	3355
OY	1397	tgtctgccttcgcgaacgatgatgtatctgttgttga	1436
Db	3359	CACCGGATTTTGGAGATACAAACCTTGTGGATGGGTGAA	3398
RESULT			
X23531			
X23531 standard; cDNA; 3045 BP.			
AC X23531;			
DE 17-JUN-1999 (first entry)			
KM Maize Xa21 gene Df4 cDNA fragment.			
KW Xa21; receptor kinase-like protein; multigene family; RKR; rice; Df4;			
OS plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.			
PN Zea mays.			
WO9909151-A2.			

QY 1330 aagaccgattcttctgtatagattatgctctttagagcttactctgacagagagct 1389
DB 1246 AAAAGCAGTGTATACCTTTGAGTGTGTGTGAGCTTATTAACAGACAGCAAGCCCA 1305
QY 1390 ttgatcttctgcctctgcagacgaltgaltgtatgtctgttgattggttaaacg 1449
DB 1306 GTAGGAATTTTGGAGAAAGAAAGAAATGACATTGTACAAATGGCGGAAAAACGGACAAAA 1365
QY 1450 ctttgaagaagaaagattgagagatgctgtcgatccctgacctgagaaacattacatt 1509
DB 1366 TGGAGCAAAAGAGGGGTGTGAAATCTTGG-----ATGAGAGCGTAAATAATGTGTGCA 1419
QY 1510 gacacagaagttgagcagctatttcaagtagcattactctgtacccaggggttcgcaatg 1569
DB 1420 ATTGTTAAAGCTATGCAAGATTTTGTGAGCAATGCTTTGTGTGAAGAGTACAGCATT 1479
QY 1570 gaggcgcttaagatgctcagagtagtccgaatgctt 1605
DB 1480 GAGAGGCGCTACAAATGAGGAAAGTAGTCCAAATGCTT 1515

RESULT 9
X23532
ID X23532 standard; DNA; 3293 BP.
AC X23532;
DE 17-JUN-1999 (first entry)
DE Tomato Xa21 clone TRK1 DNA fragment.
KW Xa21: receptor kinase-like protein; multigene family; RRK; rice; TRK1;
KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
OS Lycopersicon esculentum.
PN WO909151-A2.
PD 25-FEB-1999.
PE 17-JUN-1998; US-910386.
PR 13-AUG-1997; US-910386.
PA (NEGC) UNIV CALIFORNIA.
PI Halbert SH, Richter T, Ronald PC, Song W, Szabo V,
PI Wang G;
DR P-PSDB: W93600.
PT New RRK polynucleotides and nucleic acid constructs - used for
PT generating transgenic plants resistant to Xanthomonas
PS Claim 14; Page 59-60; 67pp; English.
CC This invention describes a method for conferring disease resistance in
CC plants. The invention describes the use of novel genes and proteins
CC belonging to the Oryza longistaminata and Oryza sativa receptor
CC kinase-like protein (RRK) Xa21 multigene family. Such genes from
CC cassava, maize and tomato are also described. The genes and proteins can
CC be used for enhancing resistance to Xanthomonas in a plant. Preferably in
CC rice or tomato.
SQ Sequence 3293 BP; 937 A; 634 C; 701 G; 994 T;

Query Match 7.7%; Score 140; DB 1; Length 3293;
Best Local Similarity 51.8%; Pred. No. 2.2e-26;
Matches 423; Conservative 0; Mismatches 375; Indels 18; Gaps 4;

QY 793 accactactgaaagaggttgatgtaaggtgataagggagccttgctgagtgcata 852
DB 2135 AACGTCATAGGAGAGAGGAGGAGGAGATGATATAGGAACTATAGCCAAATGGGAT 2194
QY 853 ctgtgagcagttaaagagcttaagaaagaaacacaggtgtgtgagctgctgcaattcaa 912
DB 2195 CATGTCCGGGTGAAGAAATTTGGAAATAGCAAAAGGCTACATGATTAACGGCCATTC 2252
QY 913 acagaggtggaagatgattgattgctgtgcatcgaaatcttctgcgtctacgtgtttc 972
DB 2252 GCTGAACCTTAACACATTAAGGAAATCAAGCATAGTACATTTGAGACATGCTCGCGTT 2311
QY 973 tgcattgacacacagagcggctcttgatatacatcatagatgagtaaggtgtgcg 1032
DB 2312 TGTTCMAACAAGAAAGTAAGTAAGTGTATAGATACATCAATGAGAGCTTGA-- 2370
QY 1033 tcgtgttaagagagcgctacgcatcagaacctcccttgattgccaactagaagag 1092

DB 2370 ----GTGAAGCTTCATAGGAAAGAGCGGCGCAATCCAAATGGAAATAGCTTAA 2425
QY 1093 attgactaggatctgctaggggcttcttattgcatgacattgatacccaagatt 1152
DB 2426 ATAGCCATGAAAGCTGCCAAGGGCCTTCTTATTTGACACAGATTCCTCCCATGATA 2485
QY 1153 atccatgtgattgtaaaagctgcgaatatattattgagcagaagaagttagggctgtgta 1212
DB 2486 ATCCACCGCATGTCATCAACAAATATATTGTAACCTGTGAACCTGMAAGCTCATGTT 2545
QY 1213 gttgatttgggttagctagctcatgattacaaaggaatacccatg--ttacaactgct 1269
DB 2546 GCAGATTTTGGATTATGCAATACTTTCGTACAAATGATAGCTTGATGCAATGTCGCA 2605
QY 1270 gtaagggttacctggtgctacatagctcccgagacctctcagctgagaaagtcacag 1329
DB 2606 ATTGCAGATCTTATGCTACATTTGCTCCAGAAATATCATACAGCTGAAATGATGAG 2665
QY 1330 aagaccgattcttctgtatagattatgctcttagagcttactctgacagagagagct 1389
DB 2666 AAAAGCAGTGTATACCTTTGAGTGTGTGTGAGCTTATTAACAGAGCAAGCCCA 2725
QY 1390 ttgatcttctgcctctgcagacgaltgaltgtatgttgttgattggttaaacg 1449
DB 2726 GTAGGAATTTTGGAGAAAGAAAGAAATGACATTTGTACAAATGGCGGAAACAGACAAAA 2785
QY 1450 ctttgaagaagaaagattgagagatgctgtgcatccctgacctgagaaacattacatt 1509
DB 2786 TGGACCAAGAAAGAGGGGTGTGAAATCTTGG-----ATGAGAGGCTAAATAATGTGTGCA 2839
QY 1510 gacacagaagttgagcagctatttcaagtagcattactctgtacccaggggttcgcaatg 1569
DB 2840 ATTGTTGAAGCTATGCAAGTATTTTGTACCAATGCTTTGTGTGAAGAGTACAGCATT 2895
QY 1570 gaggcgcttaagatgctcagagtagtccgaatgctt 1605
DB 2900 GAGAGGCGCTACAAATGAGGAAAGTAGTCCAAATGCTT 2935

RESULT 10
T62124
ID T62124 standard; cDNA to mRNA; 3176 BP.
AC T62124;
DE 17-JUN-1997 (first entry)
DE Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
KW Plant; morphogenesis; regulation; short; stem; alteration;
KW inflorescence; extraneous; gene; expression; transformation;
KW increase; control; form; length; ds.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT cds 51..2981
FT /tag= a
FT /note= "plant morphogenesis regulatory protein"
PN J09056382-A.
PD 04-MAR-1997.
PE 24-AUG-1995; 216187.
PR 24-AUG-1995; JP-216187.
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA (CHIK-) ZH CHIKUO KANKYO SANGIO GIJITSU KENKYU.
DR P-PSDB: W13408.
DR P-PSDB: W13408.
PT DNA encoding plant morphogenesis regulatory protein - useful to
PT yield plants with short stems or altered inflorescence
PS Claim 1; Pages 6-10; 17pp; Japanese.
CC The present sequence encodes an Arabidopsis thaliana plant
CC morphogenesis regulatory protein (MRP), which can be used to yield
CC a plant with, e.g. short stems or altered inflorescence. The MRP
CC acts on a plant at a specific site for a specific period, and can
CC therefore be used to regulate extraneous gene expression in a
CC plant. The MRP's cDNA or genomic DNA can be used to transform a
CC plant to increase its MRP expression, and therefore control the
CC form (particularly stem length) of the plant.

Query Match	Best Local Similarity	Matches	278;	Conservative	0;	Mismatches	211;	Indels	3;	Gaps	1;
Query Match	7.6%; Score 138.4; DB 1; Length 3176;										
Best Local Similarity	56.5%; Pred. No. 5.6e-26;										
Matches	278;	Conservative	0;	Mismatches	211;	Indels	3;	Gaps	1;		
Sequence	3176 BP;	927 A;	654 C;	659 G;	936 T;						
Query Match	7.6%; Score 138.4; DB 1; Length 3176;										
Best Local Similarity	56.5%; Pred. No. 5.6e-26;										
Matches	278;	Conservative	0;	Mismatches	211;	Indels	3;	Gaps	1;		
Sequence	3176 BP;	927 A;	654 C;	659 G;	936 T;						
Query Match	6.7%; Score 121.2; DB 1; Length 2640;										
Best Local Similarity	53.7%; Pred. No. 1.3e-21;										
Matches	321;	Conservative	0;	Mismatches	268;	Indels	9;	Gaps	3;		
Sequence	2640 BP;	800 A;	491 C;	645 G;	704 T;						

OY	791	gtacatacttggaaagagtgatgttgaagtgatataagggacgccttcgtatgct	850
Db	1630	GTACCAAACTCGGACAAAGGCGTTTGGTATTGTTTACAAAGGAGATTACTTACGCGGC	1689
OY	851	cacttctagcagtttaaaggcttaagaagaacacgaaccgggtgtgactgtcaatttc	910
Db	1690	AAGAAATCGCTGTAAAAAGGCTTTCAGAAACCTCAGTTCAAGGACGTATG--AGTTTA	1746
OY	911	aaacagaagtggaatgatattagcatgctgtgtcatcgaatcttcctgtctacgtgtt	970
Db	1747	TGATGATGAGTGACATTATATGCGCTAGGCTTCACACATATAAACCCTGTCCAAATTCCTGGCT	1806
OY	971	tcgtcatgacacacacagagcggctctctgtatcatcatcactgactgtcaatgaaagtgtg	1030
Db	1807	GTTCCATTGGAAGCAGATGAGAAACATGCTGATATATGAGATTTGGAAAATTTAAGCCTCG	1866
OY	1031	cgctgtgtttaagaagagcgctcagcctcctagaacccctcccttgattggcgaactaggaga	1090
Db	1867	ATTCTTTATCTCTTCG--GAAAAACCCAAAGGCTTAGCTTAATGTGAAGAGAGATTCG	1933
OY	1091	ggattgactiaggatcctgtcctcgaaggggcttccttaattgtcatgaccattgtatcccaaga	1150
Db	1924	ACATTACCAATGCGTGTGCTCGAGGCGCTTTTATGCTTCATCAACAGTCACAGGTTTGA	1983
OY	1151	ttatccatcgtgtatgtataaagctgcgaatatattatttggacgaagaatttgaagctgttg	1210
Db	1984	TATTCACACAGAGATTTAAAAAGTAGTAAACATTTTGCTTGATAAAAATATGATCCCAAGA	2043
OY	1211	tagtgtgatttgggtgtgcagcctcagctcatgatattacaagaataccatgtt--acaactg	1267
Db	2044	TCTGGATTTTGGGATGCGCCACAGATTTTTCGAAGGACGACGAGACTAACACATAGA	2103
OY	1268	ctctgaagggttaccttgygcatatagctcccgagfaccctcctcgaactggaagatcatcag	1327
Db	2104	AGGTGTCGGAACCTTACGCTCATCATGTCGCCAGTAGACGATGATGGATATCTCGG	2163
OY	1328	agaagacgcgatgtcttctgtgtatgggatcatgctctcttagagctctattacttgcagcag	1385
Db	2164	AAAATCAGATGTTTTCAGTTTGGAGTCATATGCTCTTGAATTTGTACTGTGNAAGAG	2221
RESULT 12			
Q32648			
ID	Q32648 standard: DNA; 2833 BP.		
AC	Q32648;		
DT	04-MAY-1993 (first entry)		
DE	S-receptor: kinase gene SRK6.		
KM	self-Incompatibility; primary; transmitter; pollen recognition;		
KM	signal; S-locus; binding domain; transmembrane; protein kinase;		
KM	glycoprotein; SUG; SRK6; SRK2; ZMPK1; ss.		
OS	Brassica oleracea.		
PF	Key	Location/Qualifiers	
FT	cds	1..2574	
FT	signal_peptide	1..97	/*tag- a
FT	misc_rna	1340..1399	/*tag- b
FT			/*tag- c
FT			/note= "transmembrane domain"
PN	EP-519869-A.		
PD	23-DEC-1992.		
PF	10-JUN-1992; 810449.		
PA	19-JUN-1991; US-717331.		
PA	(CIBA) CIBA GEIGY AG.		
PI	Nasrallah JB, Nasrallah ME, Stein J;		
DR	WPI; 92-426051/52.		
DR	P-PSDB; R29814.		
PT	New Brassica gene for S-receptor kinase, used to impart self		
PT	incompatibility to plants - has extracellular domain homologous		
PT	to S-locus glyco-protein gene		
PS	Claim 8; Page 51-52; 81pp; English.		

CC The sequences given in Q32648-49 represent two S-receptor kinase (SRK)
 CC genes. These genes are transcribed specifically in pistils and
 CC anthers and are involved in the stigmatic self-incompatibility (SI)
 CC response of Brassica. The protein products of these genes are thought
 CC to be the primary transmitters of the pollen recognition signal in the
 CC SI response. The SRK proteins comprise an S-locus binding domain, a
 CC transmembrane domain and a protein kinase domain. The S-locus binding
 CC domain is pref. substantially homologous to an S-locus glycoprotein
 CC (SLG) gene. The kinase domain has a degree of homology with the
 CC kinase domain of SRK6, SRK2, and ZMPK1.
 CC Sequence 2833 BP; 857 A; 541 C; 646 G; 789 T;

Query Match 6.4%; Score 115.4; DB 1; Length 2833;
 Best Local Similarity 53.1%; Pred. No. 4.1e-20;
 Matches 317; Conservative 0; Mismatches 271; Indels 9; Gaps 3;

OY 791 gtacatactggaagagtgattgtgtaaggtgataaagagcgccttgatgct 850
 DB 1592 GTAAACAACCTGCGACAGGTGGTATGTTATGTTTAAAGGAAGATTACTGACGGGA 1651
 OY 851 caattgtgacagttaaaggcttaagaagaacacacagcggtgtagctgagtttc 910
 DB 1652 AAGAAATTCAGTAAAGGCTATCAAGAGCTCAGTTCAGAGGAGCTGATG--AGTTTA 1708
 OY 911 aacagaggtggaatgtagcatgctgtagcagcaatctctgtagctagtggtt 970
 DB 1709 TGAAATGAGGTGACCTAATATGCGAGGCTTCACGATATTAACCTGTTAAAGTTCGGCT 1768
 OY 971 tctgcatgacacacacagcgcctctctgtatatacatagtgtaagtgtaagtggtt 1030
 DB 1769 GTTGATTTGAAGAGATGAAGATGTTGATATATGAGTATTTGAAATTTAAGCCCTG 1828
 OY 1031 cgtcgtgttaagagagcgtcagcatcagaacctccctgtagctggcacaactggaa 1050
 DB 1829 ATTCTTACT--CTATGTTAAACCCGAAAGGCTCTACCTAATGATGAGATGCG 1885
 OY 1091 gattgacactagatctgtaggggcttcttatttcagtagcacttgtagtcccaaga 1150
 DB 1886 ACATTACCAATGCTGTGCTCGAGGCTTTTATATCTTCAACAGCTCAGCGTTTAAAGA 1945
 OY 1151 ttatccatcgtgataaagcgtcaaatatataattgagaagaattgagcgtgtt 1210
 DB 1946 TATCCACAGAGATTTGAAGTAAGTAACTATTTGCTTGACAAAATATGATCCCAAGA 2005
 OY 1211 tagggatttggttagctagctcagcatgatacaagataccatgtt--acaactg 1267
 DB 2006 TCTCGATTGTTGGATTGCGCAGGATTTTGAAGGAGCAACGAAACCTTAACACAAATGA 2065
 OY 1268 ctgtaagggtagccttgtagctacagctcccgagtagctctcgactggaagtagcat 1327
 DB 2066 AGGTGCTGCGAACATACGCGCTACATGTCGCCGAAATACGCAATGATGGAATATTCGCG 2125
 OY 1328 agaagacagatgcttggttagtgatgatagtctcttagagctactacagcagaga 1384
 DB 2126 AAAAATCAGATGTTTTCAGTTTGGAGTCATAGTCTTGAATATGTTAGTGAAGA 2182

RESULT 13
 ID T13425 standard; cDNA; 2685 BP.
 AC T13425;
 DT 10-JUN-1996 (first entry)
 DE SRK-A10 cDNA.
 KW S locus receptor kinase; SRK-A10; S locus glycoprotein; SLG-A10;
 KW self-incompatibility; hybrid seed; transgenic plant; rapeseed;
 KW Brassica napus oleifera; ds.
 OS Brassica napus oleifera line W1.
 FT Key Location/Qualifiers
 FT cds 36..980
 FT misc_difference 948
 FT /tag= a
 FT /tag= b

FT /note= "there is a single base-pair deletion at
 FT position 948"

FN CA2123751-A.
 PD 12-SEP-1995.
 PF 17-MAY-1994; 123751.
 PR 11-MAR-1994; US-208909.
 PR 11-MAY-1994; US-242104.
 PA (Roth)/ Rothstein S J.
 PI Glavin TL, Goring DR, Rothstein SJ, Schafer U;
 DR WPI: 96-180298/19.
 PT DNA mols. encoding S locus glycoprotein or receptor kinase -
 PT expression vectors contg. the DNA can be used to produce
 PT self-incompatible plants, which ensure hybrid progeny and increase
 PT yield
 PS Disclosure: Fig 2; 3pp; English.
 CC cDNA clone SRK-A10 (T13425) codes for an S locus receptor kinase
 CC associated with the A10 self-incompatibility allele of Brassica
 CC napus. It was isolated from genomic DNA of rapeseed line W1 by
 CC PCR amplification. The gene contains a 1 bp deletion resulting in
 CC premature termination of translation and produ. of truncated SRK-A10
 CC protein. Expression vectors contg. a corrected SRK-A10 gene (see
 CC T13423) are used to produce self-incompatible plants able to
 CC produce hybrid seed in improved yields.
 CC Sequence 2685 BP; 816 A; 512 C; 639 G; 718 T;

Query Match 6.3%; Score 114.8; DB 1; Length 2685;
 Best Local Similarity 53.0%; Pred. No. 5.7e-20;
 Matches 317; Conservative 0; Mismatches 272; Indels 9; Gaps 3;

OY 792 taccatactggaagagtgattgtgtaaggtgataaagagcgccttgatgctc 851
 DB 1624 TAACAATAATGACACAGCGTGTGTTGGTATGTTTAAAGGAAGATTACTGACGGGCA 1683
 OY 852 actgtagaagttaaaggcttaagaagaacacacagtgtagctgagttcca 911
 DB 1684 AGAAATCGCGGTAAAGGCTATCAAAAGCTCAGTTCAAAGGCGTGTATG--AGTTAT 1740
 OY 912 aacagaggtggaatgattagatgctgtagtgaatctctgtagctacgtgtt 971
 DB 1741 GAATGAGGTGACCTTAATGCGGAGGCTTCAACATGTAATCTTGCAAAATCTTGCGCTG 1800
 OY 972 ctgtagacacacacagcgcctctctgtatatacatagtgtaagtgtaagtggttc 1031
 DB 1801 TTGCAATTGACGACGATGAGAAGATGCTGATATATGATGTTTGAAGATTAAAGCTCGA 1860
 OY 1032 gtcgtgttaagagagcgtcagcatcagaacctccctgtagtggccaactagaag 1091
 DB 1861 TTCTTATCTCTTGC--GAAAAAATCTGAAGGCTTAAGCTTAATTTGAAGAGAGATTGGA 1917
 OY 1092 gattgacactagatctgtaggggcttcttatttcagtagcacttgtagtcccaag 1151
 DB 1918 CATTAACCAATGCTGTGCTCGAGGCTTTTATATCTTCAAGACTCCCGCTTTAGAT 1977
 OY 1152 tatccatcgtgataaagcgtcaaatatataattgagaagaattgagcgtgtt 1211
 DB 1978 AATCCACAGAGATTGAAAGTAAGTAACAATTTGCTGATGAATATGTCGCCAAAGAT 2037
 OY 1212 aggtgatttggttagctagctcagcatgataaagataccatgtt--acaactgc 1268
 DB 2038 CTCGGATTTTGAATGCGCGGATATTGAAAGAGCGAGAGGAAGCTACACATGTA 2097
 OY 1269 tgaagggtagccttgtagctacatagctcccgatgacccctcgacagtggaagtagcat 1328
 DB 2098 GGTGCTGGAACTTACGGCTACATGTCGCCAGAGTACGCAATGGGTGGGATTTCTCGGA 2157
 OY 1329 gaagacagatgctctggttagtgatagtctcttagagctcactacagcagaga 1386
 DB 2158 AAATACGATGTTTTCAGTTTGGAGTCATGCTTGAATATTTACTGGAAGAGA 2215

RESULT 14
 ID T13423

ID T13423 standard; cDNA; 2686 BP.
AC T13423;
DT 10-JUN-1996 (first entry)
DE SRK-A10 gene mutant N948.
KW S locus receptor kinase; SRK-A10; S locus glycoprotein; SLG-A10;
KW self-incompatibility; hybrid seed; transgenic plant; rapeseed;
KW canola; oilseed; vector; ds.
OS Synthetic.
FH Key
FH mutation
FH location/Qualifiers
FT 948
FT /tag= a
FT /note= "a, c or g may be inserted at position 948"
PN CA123751-A.
PD 12-SEP-1995.
PE 17-MAY-1994; 123751.
PR 11-MAR-1994; US-208909.
PR 11-MAY-1994; US-242104.
PA (ROMH/) ROTHSTEIN S. U.
PI Glavin TL, Goring DR, Rothstein SJ, Schaefer U;
PI WPI; 96-180298/19.
PT DNA mols. encoding S locus glycoprotein or receptor kinase -
PT expression vectors contg. the DNA can be used to produce
PT self-incompatible plants, which ensure hybrid progeny and increase
PT yield
PS Claim 8; Fig 3; 39pp; English.
CC cDNA clone SRK-A10 (T13423) codes for an S locus receptor kinase
CC associated with the A10 self-incompatibility allele of Brassica
CC napus. The gene contains a 1 bp deletion resulting in premature
CC termination of translation and prodn. of truncated SRK-A10
CC protein. A 'corrected' gene (T13423) has a, c or g inserted at
CC position 948 (T13423). Incorporation of the corrected gene into
CC an expression vector allows the breeding of self-incompatible
CC plants able to produce hybrid seed in improved yields.
SQ Sequence 2686 BP; 816 A; 513 C; 638 G; 718 T;

Query Match 6.3%; Score 114.8; DB 1; Length 2686;
Best Local Similarity 53.0%; Pred. No. 5.7e-20;
Matches 317; Conservative 0; Mismatches 272; Indels 9; Gaps 3;

QY 792 taccactctggaagaggtgagttgtaagtggtatgaagacgcctgctgtagctc 851
DB 1625 TAAACAAATTGGACAGAGGTGTTTGTAATTGTTTACAAAGGGGAAATTACTTGACGGGCA 1684
QY 852 actctgtagcagtaaaagcttaaaagaacgaacacaggtggtgagctgcaattca 911
DB 1685 AGAAATCGGGTAAAGAGGCTATCAAGACGTCAGTTCAGAGGGGTTGAG---AGTTTAT 1741
QY 912 aacagaagtggaatgataagcagctgctgcatcgaaatctctgctgctagctgtt 971
DB 1742 GAATGAGGTGACATTAATCGCGAGGCTTCAACATGTAATCTTGCAAAATCTTGCGTG 1801
QY 972 ctgcatagaacaacaagcgcctctctatatacatcagctgaatgaaagctgtgc 1031
DB 1802 TTGCATTCGACGATGATGAAGAGTCTGATATGATGATTTTGGAAATTTAAGCTCGA 1861
QY 1032 gtccgtttaagaagagctcagccacagaacctcccctgattgccaactaggagaag 1091
DB 1862 TTCTTATCTCTCG---GAAAAACTCGAAGGCTTAAGCTTAATTTGGAAGAGATTCGA 1918
QY 1092 gattgactagatctgctaggggctctctatttgcatagcacattgtaaccagaat 1151
DB 1919 CATTACCAATGGTGTGCTCGAGGGCTTTATATCTTCATCAAGACTCCCTTAGGAT 1978
QY 1152 taccactctggtgaagaagctgcaatatattatgscgagaatttgagctgtgt 1211
DB 1979 AATCCACAGATTTGAAAGTAAGTAACATTTTCTTGCTGATAGAAATATGTCACAAAGAT 2038
QY 1212 aggtgatttggtagctagctcagctcattacaagaagctaccatgtt---acaactgc 1268
DB 2039 CTCGGATTTTGGATGCGCAGGATATTGAAAGAGCAGAGAGGAAGCTTAACCAATGA 2098
QY 1269 tctaaagggtacctgggtgctaatagctcccgagtaacctctgactggaagtaatacaga 1328

DB 2099 GGTGTCGAACTTACCGCTACATGTCCCGAGAGTACGCATATGGTGGATATTCTCGCA 2158
QY 1329 gaagcagatgctctgtggtatggtatgctcttagagctcatattacgggacagaga 1386
DB 2159 AAAATCAGATGTTTTCAGTTTGGAGTCATGTTCTTGAATATTAATCTGGAGAGA 2216
RESULT 15
T62125
ID T62125 standard; DNA; 9295 BP.
AC T62125;
DT 10-JUN-1997 (first entry)
DE Arabidopsis thaliana plant morphogenesis regulatory protein DNA.
KW Plant; morphogenesis; regulation; short; stem; alteration;
KW inflorescence; extraneous; gene; expression; transformation;
KW increase; control; form; length; ds.
OS Arabidopsis thaliana.
FH Key
FH location/Qualifiers
FH 1803.1881
FH /tag= a
FH /tag= a
FH /tag= a
FH /tag= b
FH /tag= c
FH /tag= c
FH /tag= d
FH /tag= d
FH /tag= e
FH /tag= f
FH /tag= f
FH /tag= g
FH /tag= g
FH /tag= h
FH /tag= h
FH /tag= i
FH /tag= i
FH /tag= j
FH /tag= j
FH /tag= k
FH /tag= k
FH /tag= l
FH /tag= l
FH /tag= m
FH /tag= m
FH /tag= n
FH /tag= n
FH /tag= o
FH /tag= o
FH /tag= p
FH /tag= p
FH /tag= q
FH /tag= q
FH /tag= r
FH /tag= r
FH /tag= s
FH /tag= s
FH /tag= t
FH /tag= t
FH /tag= u
FH /tag= u
FH /tag= v
FH /tag= v
FH /tag= w
FH /tag= w
FH /tag= x
FH /tag= x
FH /tag= y
FH /tag= y
FH /tag= z
FH /tag= z
FH /tag= aa
FH /tag= aa

FT Intron 4282..4349
 FT /tag- ab
 FT exon 4350..4421
 FT /tag- ac
 FT Intron 4422..4508
 FT /tag- ad
 FT exon 4509..4580
 FT /tag- ae
 FT Intron 4581..4706
 FT /tag- af
 FT exon 4707..4778
 FT /tag- ag
 FT Intron 4779..4860
 FT /tag- ah
 FT exon 4861..4932
 FT /tag- ai
 FT Intron 4933..5018
 FT /tag- aj
 FT exon 5019..5090
 FT /tag- ak
 FT Intron 5091..5176
 FT /tag- al
 FT exon 5177..5248
 FT /tag- am
 FT Intron 5249..5412
 FT /tag- an
 FT exon 5413..5481
 FT /tag- ao
 FT Intron 5482..5576
 FT /tag- ap
 FT exon 5577..5648
 FT /tag- aq
 FT Intron 5649..5726
 FT /tag- ar
 FT exon 5727..5800
 FT /tag- as
 FT Intron 5801..5882
 FT /tag- at
 FT exon 5883..6011
 FT /tag- au
 FT exon 6096..6443
 FT /tag- av
 FT Intron 6012..6095
 FT /tag- aw
 FT Intron 6444..6519
 FT /tag- ax
 FT exon 6520..5412
 FT /tag- ay
 FT Intron 6891..6974
 FT /tag- az
 FT exon 6975..7328
 FT /tag- ba
 PN J09056382-A.
 PD 04-MAR-1997.
 PE 24-AUG-1995; 216187.
 PR 24-AUG-1995; JP-216187.
 PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
 PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
 DR WPI: 97-206629/19.
 PT DNA encoding plant morphogenesis regulatory protein - useful to
 PT yield plants with short stems or altered inflorescence
 PS Claim 6; Pages 12-15; 17pp; Japanese.
 CC The present sequence encodes an Arabidopsis thaliana plant
 CC morphogenesis regulatory protein (MRP), which can be used to yield
 CC a plant with, e.g. short stems or altered inflorescence. The MRP
 CC acts on a plant at a specific site for a specific period, and can
 CC therefore be used to regulate extraneous gene expression in a
 CC plant. The MRP's cDNA or genomic DNA can be used to transform a
 CC plant to increase its MRP expression, and therefore control the
 CC form (particularly stem length) of the plant.
 SQ Sequence 9295 BP; 2873 A; 1581 C; 1602 G; 3239 T;

Query Match 6.2%; Score 112.6; DB 1; Length 9295;
 Best Local Similarity 54.5%; Pred No. 3.6e-19;
 Matches 226; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
 QY 1067 cccttgatggccaactaggaagaggtgacactagatctgtagggcttcttatt 1126
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6538 CTCTTGATGTGGACACACGGGCTTAAGATAGCATATGTGCACACAAAGTTAGCTTATC 6597
 QY 1127 tgcataccacttgatcccaagattatccatcgatgataaagctgcaatatatt 1186
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6598 TACACCATGACTGTAGTCCAGGATCATTCACAGAGACGTGAAGTGTCCAACATTCTCT 6657
 QY 1187 tggacgaagaattggaaggtctgtgtagtatttggttagctagagctcatgattaca 1246
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6658 TGGACAAAGACTTAGAGGCTCGTTTACACAGATTTTGGAAATACGAAAGCTGTGTGTGT 6717
 QY 1247 aggataccactgtacaactcgttaaggggtacactgggctacataagctccgaatacc 1306
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6718 CAAGTCACTACTTCAACTTACGTATGATGGGCAGATAGATTACATAGACCCGAGTATG 6777
 QY 1307 tctcgactggaagatcatcagaagaacggatgcttggttatgggataagctcttag 1366
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6778 CTCGCACTTCAACGGCTTCACGAGAAATCCGATGCTTACAGTTATGCAATAGCTCTTGG 6837
 QY 1367 agctcattactggacagaagagctttgatctctgccttcgaaacgataatgata 1426
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6838 AGTTGTTAACCAGAAAGCCGTTGATGACCAATCCATCCACATCTGCTTGTGT 6897
 QY 1427 tgtgttgatctgggttaaaagccttttgaagaagaagaagtggagatgctggt 1481
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 6898 CTTCCTGCTTACTCTCTCAGCTGCTGCTTGTAGGTCAAGTCCGTAAATCTTGT 6952

Search completed: June 23, 2000, 22:49:30
 Job time: 40366 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:41:28 ; Search time 274.21 Seconds
(without alignments)
860.373 Million cell updates/sec

Title: US-09-180-798-2

Perfect score: 1815

Sequence: 1 gacaataccatggaatac.....ttaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	7.7	1554	US-08-587-680A-24	Sequence 24, Appl
2	117	6.4	2571	US-07-717-331F-9	Sequence 9, Appl
3	117	6.4	2833	US-07-717-331F-1	Sequence 1, Appl
4	110.4	6.1	2749	US-08-265-628-1	Sequence 1, Appl
5	99.6	5.5	966	US-08-447-185-2	Sequence 2, Appl
6	99.6	5.5	2443	US-08-447-185-3	Sequence 3, Appl
7	92	5.1	2749	US-07-717-331F-4	Sequence 4, Appl
8	87.6	4.8	5733	US-08-473-533A-1	Sequence 1, Appl
9	58.6	3.2	3921	US-08-587-680A-3	Sequence 3, Appl
10	58.6	3.2	5992	US-08-475-891A-3	Sequence 3, Appl
11	57.2	3.2	4517	PCT-US93-06251-83	Sequence 83, Appl
12	57.2	3.0	2075	US-08-238-163-3	Sequence 3, Appl
13	54.8	3.0	6236	US-08-475-891A-1	Sequence 1, Appl
14	53.8	3.0	6256	US-08-567-375-1	Sequence 1, Appl
15	53.8	3.0	6256	US-08-587-680A-1	Sequence 1, Appl
16	53.8	3.0	2880	US-08-666-271-4	Sequence 4, Appl
17	52.6	2.9	3905	US-08-666-271-1	Sequence 1, Appl
18	52.6	2.9	2647	PCT-US93-06251-77	Sequence 77, Appl
19	52.2	2.8	2890	US-07-928-464-1	Sequence 1, Appl
20	51.4	2.8	2890	PCT-US93-07347-1	Sequence 1, Appl
21	51.4	2.8	3033	US-08-003-311B-1	Sequence 1, Appl
22	51.4	2.8	3033	US-08-261-432-1	Sequence 1, Appl
23	50.8	2.8	1804	US-08-306-691B-40	Sequence 40, Appl
24	50.8	2.8	1804	PCT-US93-06251-82	Sequence 82, Appl
25	49.8	2.7	1038	US-08-238-163-1	Sequence 1, Appl
26	49.2	2.7	831	US-08-567-375-15	Sequence 15, Appl

28	47.6	2.6	1203	US-08-602-010A-7	Sequence 7, Appl
29	47.6	2.6	1203	US-08-680-726A-7	Sequence 7, Appl
30	47.6	2.6	5495	US-08-602-010A-1	Sequence 1, Appl
31	47.6	2.6	5495	US-08-602-010A-2	Sequence 2, Appl
32	47.6	2.6	5495	US-08-680-726A-1	Sequence 1, Appl
33	47.6	2.6	5495	US-08-680-726A-2	Sequence 2, Appl
34	47.6	2.6	10592	US-08-680-726A-51	Sequence 51, Appl
35	47.6	2.6	10592	US-08-680-726A-52	Sequence 52, Appl
36	46.2	2.5	3590	US-08-547-889-1	Sequence 1, Appl
37	46.2	2.5	3590	PCT-US96-09193-1	Sequence 1, Appl
38	46	2.5	4626	US-08-306-691B-22	Sequence 22, Appl
39	46	2.5	4626	PCT-US93-06251-27	Sequence 27, Appl
40	45	2.5	7218	US-08-232-463-14	Sequence 14, Appl
41	44.6	2.5	3389	US-08-395-580-1	Sequence 1, Appl
42	44.6	2.5	3426	US-08-205-018-1	Sequence 1, Appl
43	43.8	2.4	1602	US-07-820-011A-1	Sequence 1, Appl
44	43.8	2.4	1602	PCT-US93-00445-1	Sequence 1, Appl
45	42.2	2.3	10350	US-09-047-125-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-587-680A-24
; Sequence 24, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587, 680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373, 375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475, 891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004, 645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567, 375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34, 774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA (partial)
US-08-587-680A-24

Query Match 7.7%; Score 140; DB 4; Length 1554;
Best Local Similarity 51.8%; Pred. No. 4.2e-29;
Matches 423; Conservative 0; Mismatches 375; Indels 18; Gaps 4;

```

QY 793 accaactctggaagagctggaatttgtaaggtgataaagagacgctctgcatgagctca 852
DB 715 AACGTCATAGGAGAGGAGTGAGCAAGGATAGTATATAAGGAACTATGACCAATAGGAT 774
QY 853 ctgttagcagtttaaaagccttaaaagaaacgaacacaggtgtgtgagctgagcttcaa 912
DB 775 CATGTCGGGCTGAAGAAATTTGGGATATAGCAAAAGGCTCATGATGATACGGGCTATCT 832
QY 913 acagaggtggaagatgattagatgctgtgcatgaaactctctgctctacgtgttc 972
DB 832 GCTGACTTAAACATTTAGGGAGATCAGGCATAGGTAATGATGAGACTGCTGCGCTTT 891
QY 973 tgcattgacacacacagagcagctctctgtatatacatatcagctcgttaagtgttcg 1032
DB 892 TGTTCAAACAAGAGAGTCAACTGTGCTAGTTATGAGTACATGCTAAATGGAAGCTTAG 950
QY 1033 tctgttttaaaagagcgtcagccatcagcaacctccccttgattgccaactagaagag 1092
DB 950 ---GTGAAGTCTTCATGAGGAGAGAGACGGCGCACTCCATGAGGAATGAGCTTAA 1005
QY 1093 attgacactgagatctgctcagagggctctcttatttgcatacagcattgtatcccaagat 1152
DB 1006 ATAGCCATAGAAAGCTGCCAAGGCGCTTCTTATTTGGACCAAGATGCTGCCCTATGATA 1065
QY 1153 atccatcgtatgtaaaagcctgcaaatatatttgaacgaagaatttgaagctgttcta 1212
DB 1066 ATCCACCCCGATGTCACATCCACAACTATATTGTAACCTGTGAACCTGGAAGCTCAGT 1125
QY 1213 ggtgttttgggttagcagctcagctatgatacaaggttaccatg---ttacaactgct 1269
DB 1126 GCAGATTGTGGATTAGCCAAAGTACTTGTGTAACAATGGTACTCTGTGATGTCATGTGCA 1185
QY 1270 gtaaggggttaacttgggtctatagatctccagagtaacctctcagactggaagatcagag 1329
DB 1186 ATGAGAGATCTTATGGCTATACATTGCTCCAGATATGCAATATGCACTGGAATTTGATGAG 1245
QY 1330 aagacagatgtctctgttatagtggaattatgctcttagagctcattactgagcagagagct 1389
DB 1246 AAAACCGATGTGTATAGCTTTGGAGTGGTGTGGAGCTTATTAACAGGACGAAAGCCA 1305
QY 1390 ttgtatctgctcgcctgagcgaacgatgatgttatgttattgtgttggttgaagaaagc 1449
DB 1306 GTAGCAAAATTTGGAGAAAGAAAGATGACATTTGTAACAATGGCGAAACGGAACAA 1365
QY 1450 ctttgaagagaaaagcttgagagatgctgtgcataccctgacccctgagaaacattacat 1509
DB 1366 TGGACAAAGAAAGGCGGTGTAATAATCTTGG-----ATGAGAGGCTAATAAATGTTGCA 1419
QY 1510 gacacagaagtgtgagcaactattcaagtaagcatcactctgtaccagaggttgcacatg 1569
DB 1420 ATTGTGAAGCTATGCAAGTATTTTGTGTACCAAGCTTTGTGTGAAGAGTACAGCAT 1479
QY 1570 gagcgcttaagatgtcagaggttccgaatgctt 1605
DB 1480 GAGAGGCTTACATATGAGGAAGTACTCCAAATGCTT 1515

```

RESULT 2
US-07-717-331F-9
Sequence 9, Application US/07717331F
Patent No. 5486905
GENERAL INFORMATION:

```

APPLICANT: June Nasrallah, Michael Nasrallah, and Joshua
APPLICANT: Stein
TITLE OF INVENTION: A Receptor Protein Kinase Gene
TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/717,331F
FILING DATE: June 19th 1991
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2571 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-07-717-331F-9

```

Query Match 6.4%; Score 117; DB 1; Length 2571;
Best Local Similarity 53.3%; Pred. No. 1.1e-22;
Matches 318; Conservative 0; Mismatches 270; Indels 9; Gaps 3;

```

QY 791 gtaacactctggaagaggtgatttgtaaggtgataaagagacgctctgcatgagct 850
DB 1592 GTAACAACTCGAACAAGTGGTTGGTATGTTTACAAAGGAAGTACTTACGCGGA 1651
QY 851 cactgttagcatttaagagcttaagaagaacgaacacccaggtggtgagctgagcttc 910
DB 1652 AAGAAATGCAAGTAAAGGCTATCAAGACGTCAAGTCAAGGAGACTGATG--AGTTTA 1708
QY 911 aaacagaggtggaatgattatgacatgctgtgcatcgaatctctgctctagctggt 970
DB 1709 TGAAGAGGTGACACTAATATGCGAGGCTTACGCAATATAACCTGTTCAGTTTGCT 1768
QY 971 tctgcatgacacacacagagcggctcttgatatcatcatcagatgcttaatggaagcttg 1030
DB 1769 GTTGCAATTGAAAGATGAGAGATGTTGATATATGATATTGGAATAATTAAAGCTTGG 1828
QY 1031 cgtggtttagagagaggtgacatagaacctccccttgattggtgccaactgagaga 1090
DB 1829 ATCTTATCTCTTGG---GTAAACCCGAAGGCTTAAGCTAAATTTGAATGAGAGATTGC 1885
QY 1091 ggaatgacatgagatctcgaagggcttcttatttgcatgacacattgtatcccaaga 1150
DB 1886 AACTTACCAATGTTGTGTCGAGGCTTTATATCTTCATCAACAGACACACGTTTGA 1945
QY 1151 ttatccatctgtagttaaagctcaaatatattattgagcgaagaatttgaagcgtctg 1210
DB 1946 TAATCCACAGAGATTGGAAGTAAGTAACATTGCTTGACAAAATAATGATCCCAAGA 2005
QY 1211 taggtgatttggttagtagctcagatgatttaagaagatccacatgct---acaactg 1267
DB 2006 TCTCGGATTTGGAGTGGCCAGGATATTGGAAGGAGCAAGACGAACTAACCAANTGA 2065

```


QY 1031 cgtcgtgtttaagagcgtcagccatcagaacctcccccttgatttgcgaactagaga 1090

```

; ORGANISM: Brassica napus
; STRAIN: oleifera

```

STRAIN: oleifera
INDIVIDUAL ISOLATE: W1

```

; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: S-locus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2574
; PUBLICATION INFORMATION:
; AUTHORS: GORING, DAPHNE
; AUTHORS: ROHSTEIN, STEVEN J.
; TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A
; TITLE: FUNCTIONAL SERINE/THREONINE KINASE
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749
; US-08-265-628-1

```

```

Query Match      6.1%; Score 110.4; DB 2; Length 2749;
Best Local Similarity 52.5%; Pred. No. 7.7e-21;
Matches 315; Conservative 0; Mismatches 276; Indels 9; Gaps 3;

```

```

Qy 789 tagtaaccatacttgaaagaggtgattgtaaggtgtataaagggagccttcataag 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1593 ttgttaacaaactcgacaaaggtggttcggtattggttacaaaggtatgattactatg 1652
Qy 849 ctcaacttgtagcattaaaaggttaaaagaaacagacacaggtgtagctgcaatt 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1653 gcaagaatttcgggttaaaaggtatcaaaaacgtcggttcagagagctgctg---agtt 1709
Qy 909 tcaaacagaggtggaatgattagcattgctgtagcattcgaatctctcgtctacgtg 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1710 tatgaatgagctgagattgattcgagagcttcacgattatataaacttgcgaattcttgg 1769
Qy 969 ttctcgacagcacaacaaagagcggtctctgtatcatcatatgcttaagaaagtgt 1028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1770 ctgttgcatgtagcagacgaaagatgctggtatgattgatttgaataatttaagcct 1829
Qy 1029 tgcgtcgtgttaagagagcgctcagcaccagaaacctcccttgattgycgaactagaa 1088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1830 cgattcttatctcttgc---gaaatfaaacgaagctctacgtaattgaagacagatt 1886
Qy 1089 gagggttgcaactagatctgtaggggcttcttatttgatgacattgtagatcccaa 1148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1887 caacattacaaatggtgttgcagagcctttttatcttcacaaagactcagcgtttag 1946
Qy 1149 gattatcatcgtatgtaaaagctgcacatatatttgagcagaagaatttgagctgt 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1947 gatatccacagagatataagaaagtaagtaacatttgccttgatataaaatgacaccaa 2006
Qy 1209 tgtaggtgatttggtgttagctcaggtcctcattgatacagaagataccatgtt---acaac 1265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2007 gatctcggaattttggagtgcccaaatctttgcaaggagacgaactgaactaacacaa 2066
Qy 1266 tgcgttaaggggtacctcgtgctacatagctccagtagtacctcgcagctggaagatc 1325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2067 gaagtggtcggaacttcgagctacatgctccggaatgacgaatgagatgggatttctc 2126
Qy 1326 agagaagacagatgcttctgtgtatgagatagctccttagagctcatctcgtgacagag 1385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2127 gaaaaaaacagatgttttgcagtttggagtcattgttgaattgtttagtgaaaaaag 2186

```

```

RESULT 5
US-08-447-185-2
; Sequence 2, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATHOGE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051

```

```

; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,078
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman M.F., Michael L.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)-263-1600
; TELEFAX: (716)-263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-447-185-2

```

```

Query Match      5.5%; Score 99.6; DB 1; Length 966;
Best Local Similarity 51.5%; Pred. No. 4.1e-18;
Matches 340; Conservative 0; Mismatches 299; Indels 21; Gaps 4;

```

```

Qy 752 ttctctgcgaagttcaagtcgacaggaacttttagtacc-----tacttgaa 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 ttctcttgatgatttgagagaaagcaactatatttgcatacaagttttatatttgac 145
Qy 806 gagggtgatttgtaaaaggtgtataggaagccttgctgtagtctactgttagcaggtta 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 atggctgttttggaagattttacaaaggtgtttgcctgattgagcaaaagtgccctga 205
Qy 866 aaagcttaaaagaaacagaaacacacaggtgtgtgagctgcagtttccaacagaggtgaa 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 aaagcgttaccgagcagctcctcaaaagatttgag---agttcgaaacgaataattgaga 262
Qy 926 tgattagcattgctgtcattcgaatccttctcgtctagctggttctcgtacatgaccaa 985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 ctctctcattttgcaagacatccgcatctggttcattgattgagattcttgatgaagaa 322
Qy 986 cagagcggctctctgtatcatcatatgcttaaggaagtgtgtcgtgtgttaagag 1045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 atgagatatttctatttattatataatgagaaatgggaactcctcaagacattttatag 382
Qy 1046 agcgtcagccatcgaaacctcccttgattgccaactaggaagagattgcaatagat 1105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 gatcagatcttaccacaaatgacatgagctgaggacagagctggaatgacatgagggg 442
Qy 1106 ctgctagggggtcttctatttgatgacattgtagtcccaagattatcatcagtgatg 1165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 cagccagaggtc-----TACACTTACTTCACTACTAAGCAATTAACATCGTGAAG 493
Qy 1166 taaaagctgcaatatattatgagcagaagaatttgagcgtgtgttagtatttggt 1225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 tcaacttataaactatattgcttgatgagaattttgcccataaaataattacgatttgaa 553
Qy 1226 tagctagctcattgattacaagat---accatgttacaactgctgttaggggttaact 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 tatccaaagaaagagctgagcttgatcaaacctctcttaccagatagtgcaagagacac 613
Qy 1283 tgggttagatagctcccgagtaactctcgaactggaaggtcatcagaagaagcagatgct 1342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 614 TCGGCTACATGACCCCTGATATTTTATTAAGGAGGAGCTCACTGATAAATCTGATGTTT 673
Qy 1343 ttggatagagattatgctcttagagctcattactgagacagagctttgattctgctc 1402
Db 674 ATCTTCGGTGTGTTTATTATTCAGAGTCTTGTGCTAGGCTGCGCATAGTTCATCTC 733

RESULT 6
US-08-447-185-3
; Sequence 3, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tankley, Steven D.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATHOGE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,078
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman Mr., Michael L.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1000
; TELEFAX: (716)-263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-447-185-3

Query Match 5.5%; Score 99.6; DB 1; Length 2443;
Best Local Similarity 51.5%; Pred. No. 6.7e-18;
Matches 340; Conservative 0; Mismatches 299; Indels 21; Gaps 4;

Qy 752 ttctctgagaaatgcaagtcgcaagcagatctttagtacc-----tacttgaa 805
Db 149 TTCTTTAGATGATTTGAGAGGACACATATATTTTGAATCAAGTTTATATGGAC 208
Qy 806 gaggtagattgttaaggtatagagagcctctgctgtagtcaactgttagcagtt 865
Db 209 ATGTGTCTTTGGAGAGGTTTACAGGCTGTTTGCCTGATGAGCAAGGTGCGCTGA 268
Qy 866 aaagagctaaagaagaagcaccagctgtagcgtgcaagcttcaacagaggtgnaa 925
Db 269 AAAAGCGTACCTGAGTCCCTCACAGATATGAG---AGTTCGAACAGAAATTGAGA 325
Qy 926 tgattacatggtctgcatgcaagaaatctctgctgctacgtggttctgacacaa 985
Db 326 CTCCTCATTTTGGACAGATCCGATCTGTTTCATGATAGAGATCTGTGANGAAGAA 385
Qy 986 cagagcgctctctgtatcatcatcatgctaatggaagtgtgtgctgtgttaag 1045

Db 386 ATGAGATGANTCTAATTTATTAATACATGAGATGGGAACCTCAAGAGCATTTGTATG 445
Qy 1046 agcgtacgccatcagaaacctcccttgattggcoactaggaagagattgactagat 1105
Db 446 GATCAGATCTACCCACATAGCATGAGCTGGAGCAGAGAGCTGGAGATATGATAGGG 505
Qy 1106 ctgctagggggtcttcttatttgcatgaccattgtgatacccaagattatccatcgatg 1165
Db 506 CAGCCAGAGGTC-----TACACTACCTTCATCTAGAGCAATTAATACATGAGATG 556
Qy 1166 taaagctgcaaatatatttgtagcagaagaattggagctgtgtgtagtgattgggt 1225
Db 557 TCAAGTATTAACATATGCTGTGATGAGATTTTGCCAAAAATTACGATTTGGAA 616
Qy 1226 tagctagctcatgattacaaagat--accatgttacaactgctgtaagggtacct 1282
Db 617 TATCCAGAAAGGAGCAGCTGATTATCAAAACCATCTTACACAGTGTGTAAGGAATC 676
Qy 1283 tgggtacatagctccgagtagctctgactcgactggaagtcatacagaagaacagatgtc 1342
Db 677 TCGGCTACATGACCCGTAATATTTTATTAAGGAGGAGCTCACTGAGAAATCTGATGTTT 736
Qy 1343 ttggtatgagattatgctcttagagctcaattactgagcagagagctttgattctgctc 1402
Db 737 ATCTTCGGTGTGTTTATTATTCAGAGTCTTGTGCTAGGCTGCGCATAGTTCATCTC 796

RESULT 7
US-07-717-331F-4
; Sequence 4, Application US/071717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah, Michael Nasrallah, and Joshua
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/17,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2749 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-07-717-331F-4

Query Match 5.1%; Score 92; DB 1; Length 2749;
Best Local Similarity 50.5%; Pred. No. 8.7e-16;
Matches 305; Conservative 0; Mismatches 250; Indels 9; Gaps 3;

```

QY 785 cttttagtaccacttggaaagggtgattgttgtaagggtgataaggagcgttgc 844
    || || || || || || || || || || || || || || || || || || || ||
Db 1604 CTGATTTTAAACAGAGTCGGAAGGCGTTTGGTGTGTATTACAGGGAAGTAGTGG 1663
QY 845 atggcctactgtgagcaagtaaaagcctaaagaagaacacacaggtgtgtgcgc 904
    || || || || || || || || || || || || || || || || || || || ||
Db 1664 ACGGGCAGAAATTCAGATGAGAGACTATCGGAATGTACAGTCAAGGTACGATGAGT 1723
QY 905 agttcaacaagaaggtggaatattgataagctgtgcatcgaaatcttcgtctac 964
    || || || || || || || || || || || || || || || || || || || ||
Db 1724 TCATGAACGAAGTTAGGCTATATGCAAAAGCTTACGACAAAT--AATCTGTGCGACATTC 1780
QY 965 gtgtgtctgcatagaaccaacagcgcgtctctgtatataccatcatgtaata 1024
    || || || || || || || || || || || || || || || || || || || ||
Db 1781 TTGGCTGTGTGTTTATGAGGGCGGAGAAAGATCTTAATTTACAGATCTGGAGAACTTAA 1840
QY 1025 gtgtgtcgtgtttaaagagacgltcaagcatcaaacctccctgtattgtgcaacta 1084
    || || || || || || || || || || || || || || || || || || || ||
Db 1841 GCCTCGATTCTCATCTCTTTGA--TGAAACCAAGAGCTGTATGTTAAATGGCAAAATGA 1897
QY 1085 ggaagagattgactaagactctgtcaggggccttctatttgcatagcaactgtgac 1144
    || || || || || || || || || || || || || || || || || || || ||
Db 1898 GATTGTATATCAATGATGATGCCGAGGGCTTCTATCTTACACCAAGATTCACGGT 1957
QY 1145 ccaagattatccatcgtatgataagctgcaaatatattattgagcaggaattgag 1204
    || || || || || || || || || || || || || || || || || || || ||
Db 1958 TTGAAATCATCCACAGGAGATTGAAAGCAACAAATCTTGTGTTAAAGATGACTC 2017
QY 1205 ctgtgtgaagtgatttgggttagctaggtctcattgataaagaagtaaccatgtt--a 1261
    || || || || || || || || || || || || || || || || || || || ||
Db 2018 CAATAATTTGACACTTTGGAATGGTAGAGATCTTTGAGCGGAGTAGAGCAAGCTGACA 2077
QY 1262 caactgtcgtaaaggtgactgtggtctataactccgaatgactcctgtactgtgaagt 1321
    || || || || || || || || || || || || || || || || || || || ||
Db 2078 CGAGGAAGTGTGCGGAATTAAGCTACATGCTCCAGAAATTTGCGATGAACGGACAT 2137
QY 1322 catcaagaagaacagatgctctgttatgtatgtatgtcctcctaagagcctactactgac 1381
    || || || || || || || || || || || || || || || || || || || ||
Db 2138 TCTCAATGAGTCAGATGTTTCAAGTTTGGGCTTGTGTTTGAATTAATTAAGTGGA 2197
QY 1382 agag 1385
    || ||
Db 2198 AGAG 2201

```

```

RESULT 8
US-08-473-553A-1
; Sequence 1, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
; US-08-473-553A-1

Query Match 4.8%; Score 87.6; DB 3; Length 5733;
Best Local Similarity 51.3%; Pred. No. 2.1e-14;
Matches 232; Conservative 0; Mismatches 214; Indels 6; Gaps 1;

QY 793 accatacttgaagaggtgattgtgtaagtgataaaggagcctgtcgtatgctca 852
    || || || || || || || || || || || || || || || || || || || ||
Db 4534 AACATATCCGTAAAGCCGCGAGATTTGCTACCGGTGCAATCCAAACAAACGTA 4593
QY 853 ctgttagcagttaaagggtttaaagaagaacgaacacaggtgtgagctcagttcaa 912
    || || || || || || || || || || || || || || || || || || || ||
Db 4594 GAGCTCCGATTAAAGCACTCGTGGCCGTGGACCGGAGAGGAGATGATGATTCACG 4653
QY 913 acagaggttgaagattagtcagtcgtgcacgcgaatcttcgtcgtcagttctc 972
    || || || || || || || || || || || || || || || || || || || ||
Db 4654 GCGAGATTCAAACTTTGGGGAGAAATCCCGCACCGTCAATGATGAGACTTGTGTAC 4713
QY 973 tgcattgacaccacaagagcgtctctgtatataccatcattgctcaatggaagtgcg 1032
    || || || || || || || || || || || || || || || || || || || ||
Db 4714 GTAGCCGAACAGATACGAATCTCTTTTATGAGTACATCCTTAATGAAGCTTGA 4773
QY 1033 tctgttttaagagagcgtcagccatcagaacctccctgtattgtgcacactagaaga 1092
    || || || || || || || || || || || || || || || || || || || ||
Db 4774 GAGCTTTT-----GCATGATCTAAGGTGGTCACTTCAATGGAGAGCGACATAGA 4827
QY 1093 atgcactagatctcgtcaggggcttcttatttgcatagcatcattgtatcccaagt 1152
    || || || || || || || || || || || || || || || || || || || ||
Db 4828 GTAGCCCTGGAACCTCAAGGGCTTGTGTATCTTCACACATGATGTTCAACATTTATC 4887
QY 1153 atcatcgtatgttaaagctgtcaaatatattatttgaagcaaatgttgagcgtgtga 1212
    || || || || || || || || || || || || || || || || || || || ||
Db 4888 TTGATGATGAGATTTAAGTCCAAATTAACATCTTTTGGAGCTCGATTTGAAAGCCATGTT 4947
QY 1213 gttgattgttgtagtcaggtccatgata 1244
    || || || || || || || || || || || || || || || || || || || ||
Db 4948 GCTGATTTTGGCTGTGCTAAGTTCTTAGTTGA 4979

```

```

RESULT 9
US-08-567-375-3
; Sequence 3, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Tuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:

```

```

ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/567,375
FILING DATE: 04-DEC-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645
FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058930
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product= "Xa-21"
US-08-567-375-3

```

```

Query Match          3.2%; Score 58.6; DB 4; Length 3921;
Best Local Similarity 50.1%; Pred. No. 1.6e-06;
Matches 172; Conservative 0; Mismatches 169; Indels 2; Gaps 1;

QY 86 aattctatagatcgtaaacgataataattacacagcttcacatgacct 145
DB 1126 ACTTCGCTTAGTTTCTTCTGACCTTGATTAATGAATGACAGGAGCAT 1185
QY 146 accttgacaataatgggtctctatgacatgagctttacagcaataacataagtgac 205
DB 1186 ATTGGCAATCTTATGTGGCTTACAACATC--TCTATCTCTGCAACAACAATTTCAGAGG 1243
QY 206 caattcctagatccttggaatcctgacaatttggtgagcttggaacctatataatgata 265
DB 1244 CTCCTCCATCATCGTTGGGAGGCTTAAAACTTAGGCATTCACGCTTACGAAACAA 1303
QY 266 gcttcctgacacctatacagcagacacattaggaagacctacaagctaaagcttcttcgctc 325
DB 1304 ACTTGACCGGTGCATCCCTTGCCCATAGGAATCTTACTGAACTTAATATCTTTACTGTC 1363
QY 326 tcaacaacaacagcctctcgtccaaatccaaatgctacactgataatatacaactcttc 385
DB 1364 TCGGCAACAACAATTAAGTGTGATACCATACACACTCTCAAAACCTCACAACAACTTGT 1423
QY 386 aagtcctcgattatcaacaacatcgctatcaggaccagtaacc 428
DB 1424 TGTCATTAGGCTTCAACTAATAACCTTAGTGCTCAATAC 1466

```

```

RESULT 10
US-08-587-680A-3
Sequence 3, Application US/08587680A
Patent No. 5977434
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Conferling
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,680A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645
FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/567,375
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058940US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product= "Xa-21"
US-08-587-680A-3

```

```

Query Match          3.2%; Score 58.6; DB 4; Length 3921;
Best Local Similarity 50.1%; Pred. No. 1.6e-06;
Matches 172; Conservative 0; Mismatches 169; Indels 2; Gaps 1;

QY 86 aattctatagatcgtaaacgataataattacacagcttcacatgacct 145
DB 1126 ACTTCGCTTAGTTTCTTCTGACCTTGATTAATGAATGACAGGAGCAT 1185
QY 146 accttgacaataatgggtctctatgacatgagctttacagcaataacataagtgac 205
DB 1186 ATTGGCAATCTTATGTGGCTTACAACATC--TCTATCTCTGCAACAACAATTTCAGAGG 1243

```

```

Query Match          3.2%; Score 57.2; DB 6; Length 4517;
Best Local Similarity 53.0%; Fred. No. 4.1e-06;
Matches 122; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

PCT-US93-06251-83

RESULT 12
PCT-US93-06251-83
; Sequence 83, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: WICKSTROM, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGILIO, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4517 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-83

Db 86 aattcatatgaatcgttaacagatataatatataattcaatgcgcgttcaattgacct 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1637 ACTTCCCTTAGTTGTTCTTCGACCTTGAAATGAAATGAATCACAGGAAGCATTCGAAAGAT 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 146 acctgacaacatgagggtctttagaatgagcttgaagcttacaagcttcaagcttgcgc 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1697 ATTGGCAATCTTATTGCGCTTACAAATC--TCATCTCTGCAACAACATTTTCAGAGGCT 1754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 206 caattcctagatccttgcggaaatcgtacaaatttgtagcttggaactatataatgaata 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1755 CTCTTTCATCATCGTTGGGACAGGCTTAATAAACTTAGGCATTTCTACTCGCCTCGAATAACA 1814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 266 gctctcttggaactataccggaacacattaggaaagcttacaagcttcaagcttgcgc 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1815 ACTTGACCGGTTGATCCCGTTGGCCATAGGAAATCTTACTGAACTTAATATCTTACTGC 1874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 326 tcaacaacaacacgcctctctcgttgcacaaatcgaatgcatcgaactaataatcacactctc 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1875 TCGGACACAACAATTCAGTGGTGTGATACCATACATCTCGAATCTCGAACAACCTTCT 1934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 386 aagtcctggattatcaacaacatgcgcctatagagcaagttacc 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1935 TGTCAATTAGCCCTTCACTAATAAATCTTACTGAGTGTCAATATAC 1977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

OY 1147 aagttatccatcgtgtagtgaagaagctgcgaatatatttttgacgaagaattttagact 1206
 Db 1378 AACATATTTTACCAGAGATCTTGGGCTGCTAATATTTCTTGTGGAGAAATCTTGTC 1457
 OY 1207 gcttgtagtgatttggtagtagcctaagctcatgattacaaagataccatgttacaact 1266
 Db 1438 AAAATACACAGACTTGGTTTAGCAAGSTTAATGAAGACAAATGATPACACAGACAGCAA 1497
 OY 1267 gctgttaaggaggttaccttggagctcatagctcccgagttaccttgcactgggaagtcata 1326
 Db 1498 GGTGTCAAAATTTCCAAATCAATATGGAAGACAGCTCCGGAAGCTGCACTGTATGTCGGTTTACA 1557
 OY 1337 gagagacacgcatgctcttggtagtagatcatgctcttagagctcaaac 1376
 Db 1558 ATAAAGCTGATGCTGTGCTGATTTGGAAATTTGCAACAGAACTGTGTAAC 1607

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query March	3.0%	Score 54.8	DB 1	Length 2075
Best Local Similarity	51.7%	Pred. No. 1,2e-05		
Matches 125	Conservative 0	Mismatches 117	Indels 0	Gaps 0
QY	130	ataacataagtyggaccatctccatgtgactcttgggaatctgacaattggtgagcttg	249	
Db	721	ACTATATCTACCGGAAACAAATTTCCACTCGGAATTTGGGAAGTCCACAAATCTCAAAATGTTA	780	

OY 250 gaacttaactaataagctctctcggacataccggaaacastingaaacttaacagg 309
 Db 781 AGGCTACGCTTCACCTAACCTTACAGGTCCATCCCTGAATTCCCTTAGTCAGGTGAAGAT 840
 OY 310 ctaagaattcttgcgtctcaacaacaagaagctctctgctcaattacaaatgtcact 369
 Db 841 TTGACGTTGTGTGAGTGTGAATTACAAATCAATTACCGGAACAAATCCCTTCTCCCTCT 900
 OY 370 aatattacaactctcaagtccttgattatcaaaacaatcggctatcaggaccgtacgg 429
 Db 901 CAGCTTCGGAATTTGCTATGCGATGACTTAGATCGTAAACAACCTCACCGGAACATACCG 960
 OY 430 ga 431
 Db 961 GA 962

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 3.0%; Score 53.8; DB 3; Length 6256;

[illegible]

```

15
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
10
```

```

: LOCATION: join(1648..4383, 5178..5513)
: OTHER INFORMATION: /product="RRK-F"
: OTHER INFORMATION: /note="Xa21 Xanthomonas spp. disease
: OTHER INFORMATION: resistance gene RRK-F from rice
: OTHER INFORMATION: (Oryza sativa)"
:
US-08-567-375-1

```

	Query Match	3.0%	Score 53.8;	DB 4;	Length 6256;
	Best Local Similarity	54.2%	Pred. No. 4.2e-05;		
	Matches 109;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;
Qy	179	agcttcacgaataacataagcttgagaccattccagtgcagatccttgaggaatctgcgaact	238		
Db	2885	ATCTCTGACCAACACAAATTTCAGAGGGTCATCTTCATCTCCTTGGGACGCTTGAAGACT	2944		
Qy	239	tgatgagcttgagaccatatcatgatgaatctctctgtgacctatccggaacatttggaa	298		
Db	2945	TAGGCAATTCTAGTCGGCTTCGAAACAACTTTCAGGCGTTCATCCATTGGCCATTGGAA	3004		
Qy	299	agcttacaagagcttagatctctgtgctccaaacaacacagccctctctgtgtccaatccaa	358		
Db	3005	ATCTTACTGAACTTAATATCTTACGTCTGGCACCAACAAATTCAGTGTTGGATTCACAT	3064		
Qy	359	tgctactgactaatattataca	379		
Db	3065	ACACACTCTCAACCTTCACAA	3085		

Search completed: June 23, 2000, 22:42:10
Job time: 40076 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:05:52 ; Search time 6198.48 Seconds
(without alignments)
1186.839 Million cell updates/sec

Title: US-09-180-798-2
Perfect score: 1815
Sequence: 1 gacaaataccatgtaatat.....ttaaaaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : ESR:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

No.	Score	Match	Length	DB	ID	Description
1	398.4	22.0	656	38	AA738547	AA738547 SBRK5 So
2	365.6	20.1	991	38	AA738544	AA738544 SBRK2 So
3	361	19.9	532	79	AW233982	AW233982 s132805.Y
4	352.6	19.4	555	51	AI728030	AI728030 BNLGH1962
5	349	19.2	1435	38	AA738545	AA738545 SBRK3 So
6	332	18.3	538	62	AI900110	AI900110 sc01a04.Y
7	279.8	15.4	555	62	AI896277	AI896277 EST265720
8	274.6	15.1	597	51	AI727837	AI727837 BNLGH1920
9	268.2	14.8	655	46	AA738546	AA738546 SBRK4 So
10	250.4	13.8	703	51	AI729440	AI729440 BNLGH1133
11	220.2	12.3	471	47	AI486453	AI486453 EST244774
12	220.6	12.2	380	43	AI165643	AI165643 A087F06u
13	201.2	11.1	412	79	AW310334	AW310334 s434F03.x
14	198.4	10.9	474	46	AI416897	AI416897 sal9611.x
15	186.6	10.7	524	60	AW033961	AW033961 EST277623
16	194.8	10.8	475	80	AW329855	AW329855 N201132e
17	181.4	10.0	299	20	D25047	D25047 R1CR2976A.R
18	180.2	9.9	714	92	AQ069360	AQ069360 LERUM53TR
19	178.6	9.8	490	47	AI486325	AI486325 EST244646
20	172.6	9.5	1202	84	B09168	B09168 T2E10-T7 TA
21	159	8.8	1434	63	AI967314	AI967314 U11mpes
22	158.6	8.7	383	81	AW423485	AW423485 sh67e01.Y
23	158.4	8.7	631	64	AW067518	AW067518 660013E08
24	149.8	8.3	647	51	AI730535	AI730535 BNLGH1700
25	145.8	8.0	621	92	AQ069361	AQ069361 LERUM53TR
26	144.4	8.0	550	74	AW223870	AW223870 EST300681
27	139.8	7.7	548	51	AI728857	AI728857 BNLGH1118
28	139	7.7	608	79	AW234133	AW234133 s134F03.Y
29	138	7.6	496	62	CNS007F9	AL096099 Arabidops
30	134.2	7.4	496	63	AI996242	AI996242 701550775
31	133.2	7.3	385	49	AU056335	AU056335 AU056335
32	133.4	7.3	717	49	AU056601	AU056601 AU056601
33	129.4	7.1	536	62	AI896183	AI896183 EST265626
34	128	7.1	428	28	AA080670	AA080670 EST012 Su
35	125.8	6.9	636	23	T45217	T45217 8480 Lambda
36	125.4	6.9	572	80	AW284352	AW284352 LG1.275.D
37	125	6.9	1345	64	AI967315	AI967315 U11mpes
38	124.8	6.9	550	64	AW042251	AW042251 614026B11
39	124.2	6.8	747	49	AO840411	AO840411 nbx0058F
40	123.2	6.8	477	74	AW224241	AW224241 EST300968
41	122.6	6.8	304	35	C22490	C22490 C22490 Rice
42	122.4	6.7	679	63	AI937984	AI937984 sc08e07.Y
43	120.6	6.7	663	74	AW201125	AW201125 se98B06.Y
44	120.8	6.7	753	105	AO271213	AO271213 nbx00251
45	119.8	6.6	606	51	AI729170	AI729170 BNLGH1128

ALIGNMENTS

```
RESULT 1
AA738547 656 bp mRNA EST 17-AUG-1998
DEFINITION SBRK5 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
            SBRK5 similar to protein kinase, highest similarity to
            receptor-like protein kinases, mRNA sequence.
ACCESSION  AA738547
VERSION     AA738547.1 GI:3421472
KEYWORDS    EST.
SOURCE      sorghum.
ORGANISM    Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; Sorghum.
REFERENCE   1 (bases 1 to 656)
            Annen, F., Chang, J.-D., Paterson, A.H. and Stockhaus, J.
            Characterization of 14 different putative protein kinase cDNA
            clones of the C4 plant Sorghum bicolor
            Mol. Gen. Genet. 259 (1), 115-122 (1998)
JOURNAL     98409267
COMMENT     On Nov 29, 1993 this sequence version replaced gi:636097.
```

```

Contact: Stockhaus J
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49) -211-81-14953
Fax: (49) -211-81-14871
Email: stockha@uni-duesseldorf.de

Insert Length: 1100 Std Error: 0.00.
Location/Qualifiers
1..656
/organism="Sorghum bicolor"
/strain="cv. TX430"
/db_xref="taxon:4558"
/clone="SBRK5"
/clone_lib="Sorghum bicolor cv. TX430 leaf"
/dev_stage="green"
/note="Organ: leaf; Vector: Lambda ZAP II; Site_1: Eco RI;
Site_2: Xho I; leaf cDNA library from green leaves,
unidirectionally cloned"

BASE COUNT  161 a 145 c 184 g 166 t

ORIGIN
Query Match 22.0%; Score 398.4; DB 38; Length 656;
Best Local Similarity 76.5%; Pred. No. 5.3e-84;
Matches 502; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

QY 961 ctacgtggttctcgtacgacacacagagcggtcttctgtatccatcattgctat 1020
DB 1 CTTCTGGTTCCTGCTGACGCGCTACTGACCGTGTAGTACCATGAGTACTGCTACT 60

QY 1021 ggaagtggtgctggttaagagagcgctgacatgaacccctccctgattgg--- 1078
DB 61 GGAAGTGTGACATCAGCTTACGAGAGCAGACGACATGACACCTGTTAAGTGGAA 120

QY 1078 ccaacaggaagagagctgacatgacatgctgtaaggggcttctctatttgcattgacat 1137
DB 121 ACACAGAGAGAGAGAGATGCTGCGTGGATGTGCAAGAGGACTTCTTACGTGATGATC 180

QY 1138 tgtgatcccaagattatccatcgtgatgtaaaagctgcaatatatttggagagaa 1197
DB 181 TGCATGCCCAAAATCATCATCGGATCGGATCCAAAGCTGCAAAATTCCTTTCATAGAC 240

QY 1198 ttgagagcgtgtgattgatttgggttagctagctgacatgattatgaagataccat 1257
DB 241 TTCAGAGCAGATGTGTGGGTATTTTGGGCTTCCAGCTTAATGACTACAAAGATCCCAT 300

QY 1258 gttaacactgctgtaaggggtacctggtctacatagctccagttacctcgactgga 1317
DB 301 GTCACAACTGCTGCTCGTGGAAACAAATGAGACAAATGGCCCTGAGTACTATCCACTGGC 360

QY 1318 aagtcacagagaagacgcgattgcttgggtatgggattatgctttagagcttact 1377
DB 361 AATGCTCTGAAAAGACTGATGTTTGGCTATGGATGATCATCTCTGAGATTTTACT 420

QY 1378 ggaacagagcctttagattgctgctgctgcaacagatgagatgattatggttggat 1437
DB 421 GGCACAGGGCATTTATATCTTGGCTGCTCTTGCATAATGATGAGAGATTTATGCTTTGAC 480

QY 1438 tgggttaaaagcctttagaagaaagttggaatgctggttcgactcctgactggag 1497
DB 481 TGGGTGAAGGATGCTGTAAGAGGAAGAGGAGGAGATCTGCTGGACCAATCTGGAG 540

QY 1498 aacaattacattgacacagaagtgtgagcagcttatccaagtgcattactgtgaccag 1557
DB 541 AAGGCTTACAGAGATTCAGAGGTGAGGAACCTGATCCAGCTGCTCTGCAAGCAG 600

QY 1558 ggttcgccaatgagagcgcttaagatgctcagaggttagtcggaattcttgaaggtga 1613
DB 601 GGCTCCCGCTGTGACCTCCCAAGATGTCGAGGTGAGGATGCTGCAAGGCTGA 656
```

```

RESULT 2
LOCUS   AA738544      991 bp      mRNA      EST      17-AUG-1998
DEFINITION
SBRLK2 Sorghum bicolor cv. TX430 leaf sorghum bicolor cDNA clone
SBRLK2 similar to protein kinase, highest similarity to
receptor-like protein kinases, mRNA sequence.
ACCESSION
VERSION  AA738544.1  GI:3421469
SOURCE   EST.
ORGANISM sorghum.
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
           Poaceae; Sorghum.
REFERENCE
AUTHORS  Annen, F., Chang, J.-L., Paterson, A.H. and Stockhaus, J.
TITLE     Characterization of 14 different putative protein kinase cDNA
JOURNAL  Mol. Gen. Genet. 259 (1), 115-122 (1998)
MEDLINE  98409267
COMMENT   On Nov 29, 1993 this sequence version replaced gi:636080.
Contact: Stockhaus J
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49)-211-81-14953
Fax: (49)-211-81-14871.
Email: stockhaus@uni-duesseldorf.de

FEATURES
source      Insert Length: 1500 Std Error: 0.00.
            Location/Qualifiers
            1..991
            /organism="Sorghum bicolor"
            /strain="cv. TX430"
            /db_xref="taxon:4558"
            /clone="SBRLK2"
            /clone_lib="Sorghum bicolor cv. TX430 leaf"
            /dev_stage="green"
            /note="Organ: leaf; Vector: lambda ZAP II; Site.1: Eco RI;
            Site.2: Xho I; leaf cDNA library from green leaves,
            unidirectionally cloned"

BASE COUNT      253 a      194 c      293 g      251 t
ORIGIN
Query Match      20.1%; Score 365.6; DB 38; Length 991;
Best Local Similarity 64.4%; Fred. No. 3.1e-76;
Matches 599; Conservative 0; Mismatches 319; Indels 12; Gaps 3;

QY 734 ttgttcaactgaagaggtttctctcgcagaaatgcaagtcgcaacgatactttagta 793
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 17 TTGGCCAGTTGATAAAGATTTCAGAGAGAGATTGCACACTTGCACAGTAATTTTCAGTG 76
QY 794 ccata-----cttggaagaggttgatctgtgaagtgataaaggagcgccttgctgaty 847
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 77 AGAAATAATGTTCTTGACAGCAAGGGGCTTTGGGAAAGTAATTAAGGACACTTCCAGATG 136
QY 848 gctcaactgtgacaggttaaaaggcttaaaagaagaacgaacaggtgtgagctgacat 907
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 137 GCATTAAGATTGCTGTAAACGGTTAACTGATTTTCGAAAGTCTTGAGGAGAGCTGCTT 196
QY 908 ttcaaacagaggttgaaatgattagcatggtcgtgcatcgaatctctcgtctcagtg 967
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 197 TCTTGCGAGAGGTGAGAGCTGATAGTGTGAGATTCACCGGAATCTTTAAGATGATG 256
QY 968 gttctgcatgacacacacagagagcgctctctgtatattccatcatcgtgcaatggaagtg 1027
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 257 GTTCTGTACACACAAACAGAGCGCTGCTGTTTATCTTTCATGACACAACTTATG 316
QY 1028 ttgcgtcgtgttaagaagcgtcagcatcgaacacctcccttgattgccaactagg 1087
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 317 TT---TCTACCGTCTACAGGAATTTAACCTGGGAGGACATATTGATGTCTGCAGCA 373
QY 1088 agaggattgcactagatcctcgtcaggggccttcttatttcatacgtacacatgtatccca 1147
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 374 AGCAGAGGCTAATAGCAGACAGCTCGTGAGACTGAGATTTTGCACAGACATGCAATCTTA 433
QY 1148 agattatccctcgttgatgaaagctgcaaatatattatgtgacgaagaattgagtg 1207
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 434 AGATTATACATCGTGAACGTCAAGCTGCGCAATGCTTGCTTATGATAGGTTTGAACGG 493
QY 1208 ttgtagtgatttttgggttagcgtcgaatgattacacagaatacccatgtacaactg 1267
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 494 TTGTGTGATTTTCGGCTTGCGCCACACTGCTGGATGATACAGAAAGCACTGACTACTAC 553
QY 1268 ctgtaaaggggtacctgtgggtacatagctcccgagtaacctctcgactggaaatcatag 1327
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 554 AGGTCCGTGGAACAAATGGGTGACATTTGCCCGGAAATTTGCCACTGGGAAGTATCCG 613
QY 1328 agaagacagatgcttctgttgatgattatgctcttaagagctcaattacagacagag 1387
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 614 AGAGAACCGATGTTTGTGTTGCTTACGCAATATGCTTCTGAGCTAGTCACTGTCAGCGAG 673
QY 1388 ctttgatctgtcgtcgcttcgcagacatgatgatgtatgtgtgtgattggtttaa 1447
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 674 CCATGACACTTTGACGCTGAGAGAGAAAGATGATG---TGTACTTGTATCATGTCAACA 730
QY 1448 gcccttgaagaagaaagtgtgagatgctggtgatcctcgaacctggagaaatattca 1507
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 731 ACGTGCAAAAGGAGGACATCTGACGCCCATGATGACCGTAACGTAACGAACAATTTACA 790
QY 1508 ttgacacagaagttagagcaagcttatcaagtagaattactctgtacccagggttgccaa 1567
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 791 ACGGCGAGAGAGGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850
QY 1568 ttgagcggtcctaagaatgtagaggttagtcgaatgctcttgaagtgatgcttcgaaga 1627
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 851 AGGACCGCGCGTCCATCTCCAGAGTGGTCCGATGCTGAGAGCGAGGCTTCGACAGAGA 910
QY 1628 agtggagcagatggcaaaagttagaagta 1657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 911 GGTGGAGAGAGTGGCAGCAGAGTGGAGGTGA 940

RESULT 3
LOCUS   AM233982      532 bp      mRNA      EST      13-DEC-1999
DEFINITION
sf32g05.v1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl02e-1713 5' similar to TR:023921 023921 SOMATIC EMBRYOGENESIS
RECEPTOR-LIKE KINASE.; mRNA sequence.
ACCESSION
VERSION  AM233982
KEYWORDS  AM233982.1 GI:6566309
SOURCE   EST.
ORGANISM soybean.
           Glycine max
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
           eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
           Glycine.
REFERENCE
AUTHORS  1 (bases 1 to 532)
           Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
           Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
           Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,
           Bowers, Y., Person, B., Swaller, T., Glendon, M., Pape, D., Harvey, N.,
           Schum, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
           McCann, R., Waterson, R. and Wilson, R.
           Public Soybean EST project
           Unpublished (1999)
TITLE     On Feb 18, 1993 this sequence version replaced gi:4297707.
JOURNAL  Contact: Shoemaker R/public Soybean EST project
COMMENT   Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
           Tel: 314 286 1800

```

Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 403.

FEATURES

SOURCE

1. 532
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1713"
/clone_lib="Gm-cl028"
/tissue_type="roots of 'superpod' plants"
/lab_host="DH10B"

/note="Vector: Bluescript II XR, Site-1: EcoRI, Site-2: XhoI; The mRNA was isolated from roots of Glycine max 'superpod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A-C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGACTACTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI: all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Kelm and Dr. Virginia Coreyell."

BASE COUNT 145 a 92 c 129 g 165 t 1 others
ORIGIN

Query Match

Best Local Similarity 19.9%; Score 361; DB 79; Length 532;
Matches 424; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 943 catcgaaatctctcgctcgaatggttctgcatgacacacagagcgtctcttcta 1002
DB 2 CATCGTAATTTCCTGGCTTCGTGTTTGTATGACACTTACGAGCGTGGTTGTG 61
QY 1003 tatccatcatgagcgaatggaagtgttcgtcgttttaagaagagcgtccacatcaga 1062
DB 62 TATCTTACATGCTACGAGAGTGTACATCATTTTACGAGACGTCACAGAAATCCCA 121
QY 1063 cctcccttgatctgccaactgaagagatgactgaagatctgctgaaggggtctctc 1122
DB 122 CCGCCACTTGGCTGCGCAACGAGACGATTCATTCGATTCGCAAGGGGCTTCT 181
QY 1123 tttttgacgaccatttgatcccaagatcatcctgtagatgaagcctgcaaatata 1182
DB 182 TATTTGCATGATCATTTGAGACCTTAAGATTATTCACCGGATGCTAAAGCACCTAATA 241
QY 1183 ttattgagaaagattgaagcgtctgtagtgatttggttgtagcgaagctcatgat 1242
DB 242 TTGTTGAGAGAGAAATTGTAACAGACTTGTGAGATTTTGGTTTACCTTAAGATTGAT 301

QY 1243 tacaagataccatgattacacactgctgtaaggggtacattggtgctacatagctcccgag 1302
DB 302 TATTAAGATACATCATGTTACTACTGCTGTAGCTGAGAAATTTGGACATATGACACAGCA 361
QY 1303 tacctctgactgtaagatgcatcagaagaaccgagctgcttgattgattgattatgct 1362
DB 362 TACCTCTCAACTGGAAGAGCTTCAGAGAGACTGATGTTTGGATATGATGATGCTT 421
QY 1363 tttagagctcatcactgagcagagagcttttattgactgctgccttgagcgaagatgat 1422
DB 422 CTGGAAGTATATACGTGGAGAAAGGCTTTGATCTACTGCACTTGCACATGATGAT 481
QY 1423 gtatattgcttgattgattgattgattgattgattgattgattgattgattgattgatt 1472
DB 482 GTCATGTTCTGATTTGGTTTGAAGAGCTTCACAGAGACAGAAATGGA 531

RESULT 4

LOCUS A1728030 555 bp mRNA EST 11-JUN-1999
DEFINITION BNGH19629 SIX-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase [Daucus carota], mRNA sequence.

ACCESSION A1728030
VERSION A1728030.1 GI:5046816
KEYWORDS EST.

SOURCE upland cotton.
ORGANISM Gossypium hirsutum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 555)
AUTHORS Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.

TITLE ESTs from developing cotton fiber
JOURNAL Unpublished (1999)
COMMENT On Jun 5, 1998 this sequence version replaced gi:3187310.

CONTACT: Ben Burr
Biology Department
Brookhaven National Laboratory
Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
Email: burr@bnl1.bnl.gov
Seq primer: T3 Primer.
Location/Qualifiers

FEATURES

source

1. 555
/organism="Gossypium hirsutum"
/cultivar="Acacia Maxxa"
/db_xref="taxon:3635"
/clone_lib="Six-day Cotton fiber"
/tissue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="XLI-Blue"
/note="Vector: pBluescript II KS+"

BASE COUNT 146 a 100 c 147 g 162 t
ORIGIN

Query Match

Best Local Similarity 19.4%; Score 352.6; DB 51; Length 555;
Matches 427; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1205 ctgttgtagtgatttggttgtagctagctcatgattacaaagatataccatgtaaca 1264
DB 1 CTGTTGTGTTGACTTTGGTTGCTTAACCTATGATGACTACAGATACCATGTAACTA 60
QY 1265 ctgctgtaaggggtacactgggtctacatagctccgagtaactctcgatggaagatc 1324
DB 61 CTGCTGTAAGCTGACAAATTTGACATATTCCTCTGATATCTCTACTGGAATAATCTT 120
QY 1325 cagagaagaccagatgcttggtgattgattgattgattgattgattgattgattgattgatt 1384
DB 121 CAGGAAATCTGATGTTTGGGTATGATCATGCTTTTGGAGCTTATTAATGAGACAC 180

QY	1385	gaattttgatctctgctcccttgcgaacagtaatagtttatggttggttggtta	1444
Db	181	GGGCTTTGATCTGCTGCTCTTGCAAAATGATGATGTCATGTTGCTTGATGGGTCA	240
QY	1445	aaagcccttttgaagaagaaaaagtttggagatgcgtgcgcactccagaccctggagaacaatt	1504
Db	241	AAGGACTTCTGAAAGGAGGAAAAAGCTGGAAATTCGTAATGCTGATCTGSCAAACCAANT	300
QY	1505	acattgcacacgaagaagtttgtagcagcttatttcataagtaactactctgtaccaggagttcgc	1564
Db	301	ATGTAAAGAACTGAGGTAAAGCAGTTAAACCGAGTTCCTGCTATGCACACAAGGTTCC	360
QY	1565	caatgtagagggccttaagatgctaaaggtgaagccgaatgcttgaagtgatgagccttgcaag	1624
Db	361	CAATGGACCGGGCCAAAGATGTGAGAAAGTGTTAAATGCTGGAAGGTGATGGTTGGCCG	420
QY	1625	aaaagtggagacgaatgtgcacaaagttagaagtcacatccataagaacgttagaattagctccac	1684
Db	421	AAAGATGGGATGATGTCGGCAAAAGTTGAAGTTCACGGCAGAGGTTGAACTTGGCCCCC	480
QY	1685	atcgaaccttcgaatggaatcctagaccctgcacagataaacttgatgcatgcttttgaattatctg	1744
Db	481	ATCCGTAATTCGTGATTGGATTGTGGACCTCAACTGACAAATTCGACATGCTGTTAGATTTCGG	540
QY	1745	gtcccaagataa	1755
Db	541	gtcccaagtgta	551

RESULT	5
AA738545	
LOCUS	
DEFINITION	
AA738545	1435 bp mRNA EST 17-AUG-1998
SBDRK3 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone	
SBDRK3 similar to protein kinase, highest similarity to	
receptor-like protein kinases, mRNA sequence.	

ACCESSION	AA738545
VERSION	AA738545.1
	GI:3421470

KEYWORDS

SOURCE ORGANISM

REFERENCE

AUTHORS

TITLE

ТОПКАТ

JOURNAL
OF
MEDICAL
NUTRITION

COMMENT

APPENDIX

FEATURES

2000

BASE COUNT

Query Match	19.28;	Score 349;	DB 28;	Length 1435;
Best Local Similarity	64.5%;	Pred. NO. 2.4e-17;		
Matches 555;	Conservative 0;	Mismatches 300;	Indels 6;	Gaps 2;
Qy 797	tactgtgaagaaggttgattgtgttaagtgatataagaagacgcctgtcgtatgctcacttg	856		
Dd 530	TTCTGTGACACAGGGGGCTTTGGAAAGTATATAAAGAGACATCTCCAGATGGCACTTAAG	589		
Qy 857	tagcagttaaaggtcgttaagaagaacgaacacacagtgtygtagctcagtttcaaacag	916		
Dd 590	TTGCTGTAAACCGTTAACTGATTCAGAAAGCCCTGGTGAGAGGCGCTTTTTCGCTG	649		
Qy 917	agctggaataagattatgaatcagctgcatctgaaatctctcgcctcagtggtttctga	976		
Dd 650	AGCTTAGCTGATTATAGTTCAGTTCACCGAATCTTTTAAGATTGATTTGTTCTGTA	709		
Qy 977	tgaacacaacaagacgagctcttctgtatatacatalatgcatatgtaagaaagtgtgcgt	1036		
Dd 710	CAACACAAACAGAGCCCTGCTGTTATTCCTTCATCCAGAAATCTTAAGTGGTGGCTTACC	769		
Qy 1037	gttaaaagcgtcagccatcagaacccccccttgattgtgccaactagaagaagatg	1096		
Dd 770	GCTTACAGCA---ATTAACTGGAGACCAATATTAGATTGTCTGCAAGAGACGGAGTG	826		
Qy 1097	cactagagatcgtctaaagggtcttctatcttcatgcatgacaactgtgatacccaagattacc	1156		

[illegible]

RESULT	6
AIR00110	
LOCUS	AIR00110
DEFINITION	sclotad.y1 Gm-cl012 bp cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl012-871 5' similar to TR:O23921 O23921 SOMATIC EMBRYOGENESIS

RECEPTOR-LIKE KINASE. ; mRNA sequence.
 ACCESSION AI900110
 VERSION AI900110.1 GI:5606012
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 1 (bases 1 to 538)
 REFERENCE Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT On Jun 5, 1998 this sequence version replaced gi:3189472.
 Contact: Shoemaker R./Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 411.
 Location/Qualifiers
 1. 538
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: GM-cl012-871"
 /clone_1lb="GM-cl012"
 /tissue_type="Apical shoot tips, 9-10 day old etiolated seedlings"
 /lab_host="XLI-Gold"
 /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from the apical shoots of 9 to 10 day old etiolated seedlings. The shoot tips including any emerged leaves were harvested for mRNA isolation. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 153 a 105 c 142 g 138 t
 ORIGIN

Query Match 18.3%; Score 332; DB 62; Length 538;
 Matches 401; Conservativity 0; Mismatches 115; Indels 0; Gaps 0;

QY 1242 ttacaagatacccatgttacaactgtctaagggtactccttggtacataagtcacca 1301
 DB 1 TTACAAGACACGCATGTGACAACTGCTACGGGGCCTATCGGGCATATAGCTCTGA 60
 QY 1302 gtactcgcactggaaggtatcagagaagcagtgcttggtatgagatagct 1361
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 61 GTACCTTCCACTGGTAAATCTTCAGAGAAACTGACGTTTTGGTTATGATATCATGCT 120
 QY 1362 cttagagcttactttagagacagagagcttttagcttgccttcgcgaacagatgata 1421
 DB 121 TCTGAGACTGATACCTGACAGAGAGAGCTTTGACCTGCTGCATCGCTAAGACATGA 180
 QY 1422 tgtatgtcttgtagatgggttaaaagccctttgaaagagaagatgtgagatgct 1481
 DB 181 TGTATGCTCTGATGATGGGTAAAGAGCTTCGAAAGAGAAAAAAGCTTGAATGTGGT 240
 QY 1482 cgaaccctgagagacattacattgacacagaagctgagcagcttaccagtagc 1541
 DB 241 AGATCCGATCTACAAACCACTATATAGAACTGAGTAGAACAATATACAGGTTGC 300
 QY 1542 attactgttaccagaggttcgccaatgagagcgccctagaatgctagagtgctcgaat 1601
 DB 301 ACCTACTGACACACAAAGATGCCCGATGAGCCACCACTAGATATCAGAAAGTGTGAAT 360
 QY 1602 gcttgaggtgtagtgccttcacagaaagtggagcaggtgcaaaagttagatcaca 1661
 DB 361 GCTTGAAGGTGATGCTTGGCAGAAAGATGGATGAGTGCAAAAGTGGAATTTCCG 420
 QY 1662 tcaagacgtagaattagctccacatcgaactctggaatgcttagactcgacagataa 1721
 DB 421 GCAGAGAGTGAGCTGGCCCTCATCTTAATTCGATGATGATGCTGACACCGAAAA 480
 QY 1722 ctgcatgctttgtaattatcttggtccaagataac 1757
 DB 481 TCTGATGCACTAGATGATTCTGTGTCAGAGTGACC 516

RESULT 7
 A1896277 555 bp mRNA EST 27-JUL-1999
 LOCUS A1896277
 DEFINITION EST265720 tomato callus, TMU Lycopersicon esculentum cDNA clone
 ACCESSION C18C14J5, mRNA sequence.
 VERSION A1896277
 KEYWORDS A1896277.1 GI:5602179
 SOURCE EST.
 ORGANISM tomato.
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 555)
 REFERENCE Alcala, J., Vrebalov, J., White, R., Matern, A.L., Visions, T., Holt, L.E., Liang, F., Opton, J., Craven, M.B., Bowman, C.L., Ahn, S., Roming, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 TITLE Generation of ESTs from tomato callus tissue
 JOURNAL Unpublished (1999)
 COMMENT On May 18, 1998 this sequence version replaced gi:3137451.
 Contact: David Frisch
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@CLEMSON.EDU
 5 prime sequence.

FEATURES
 source
 1. 555
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="C18C14J5"
 /clone_1lb="tomato callus, TMU"
 /tissue_type="callus"
 /dev_stage="25-40 days old"
 /lab_host="XLI-Blue MRF"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni laboratory; c18C - Cytiledons


```

VERSION AA738546.1 GI:4242614
KEYWORDS EST.
SOURCE Sorghum.
ORGANISM Sorghum bicolor
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
          Poaceae; Sorghum.
REFERENCE 1 (bases 1 to 655)
AUTHORS Anen, F., Chang, J.-L., Paterson, A.H. and Stockhaus, J.
TITLE Characterization of 14 different putative protein kinase cDNA
        Clones of the C4 plant Sorghum bicolor
JOURNAL Mol. Gen. Genet. 259 (1), 115-122 (1998)
MEDLINE 98409267
COMMENT On Jan 13, 1998 this sequence version replaced gi:3421471.
        Contact: Stockhaus J
        Institut fuer Entwicklungsbiologie und Molekularbiologie der
        Pflanzen
        Heinrich-Heine-Universitaet
        Universitaetsstrasse 1, 40225 Duesseldorf, Germany
        Tel: (49)-211-81-14953
        Fax: (49)-211-81-14871
        Email: stockhaus@uni-duesseldorf.de

FEATURES
    source          Insert Length: 600 Std Error: 0.00.
                    Location/Qualifiers
                        1..655
                        /organism="Sorghum bicolor"
                        /strain="cv. TX430"
                        /db_xref="taxon:4558"
                        /clone="SDRLK4"
                        /clone_id="Sorghum bicolor cv. TX430 leaf"
                        /dev_stage="green"
                        /note="Organ: leaf; Vector: Lambda ZAP II; Site 1: Eco RI;
                        Site 2: Xho I; leaf cDNA library from green leaves,
                        unidirectionally cloned"

BASE COUNT      169 a      124 c      179 g      183 t
ORIGIN
Query Match          14.8%; Score 268.2; DB 46; Length 655;
Best Local Similarity 65.3%; Pred. No. 3.1e-53;
Matches 412; Conservative 0; Mismatches 213; Indels 6; Gaps 1;

744 gaagaggtctctcgcgagaattcgaagtcgaagctactttagt-----accat 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 GAAAGATTTCGATGCGAGATTTCGAACTTCGAACTGATATATTTCAGTGAGAAAATGT 62

798 acttggaaaggttggaattgttaagggtatagaaggacgctctgtagtgcactgt 857
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 TCTTGGAAGAGGGGCTTGGGAAAGTATATAAGGAGCAGCTCCAGATGCGACTAAGAT 122

858 agcagtttaaaggcttaaagaagaacacacccagggtgtgaagctgcaggttcaacaga 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 TGCTTTAAACGCTTAACCTGATTACGAAAGTCCTGCTGAGAGGCTGCTTCTTGCGTGA 182

918 ggttgaatatgattagcagctgtgtcatcgaacatctctcgtcgtcagctgttctgcac 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GGTGAGCTGATTAGTGTTCGAGTTCACCGGAGATCTTTAGATTGATGGTTTCTGAC 242

978 gaacacaagaagggctctctgtatcatcacatcagtgctaaatggaagtgttcgtcgtg 1037
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 AACACAACAGAGCCCTGCTTGTATTACCTTTCAATGACAGATCTTAAGTGGCCTACCG 302

1038 tttaagagaagcgtcagcatcagaacccccccttgatgtggccaactggaaagagattgc 1097
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 TCTACGAGATTATAACCTGGGGACCAATATTAGATTGGCTCGAAGGAAGCAAGTGC 362

1098 actaagatcgtcgaagggtcttcttattcgaatcgaatcgtgagttcccaagattatcca 1157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 TATAGGCACAGCTCGTGGAGCTGAGTATATGACAGACAGACGCAATCTTAAGATTATACA 422

1158 tcgtgatctaaagctgcgaatatattatgtgaacgaagaattggaagctgtgtgaagtga 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 423 TCGTAGCTCAAGAGCTGCCAATGCTCTGCTCATGAGATTGTAACCGGTTGTGTGA 482
QY 1218 tttaggttagctagctcatatgattacaagatgccatgttacaattgctgaagg 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 TTTCCGGCTTGGCCAAAGCTGTGGATGTACAGAAACATCTGTGACTACTCAGGTCGTGG 542

1278 taccttggctacatagctcccgagtaacctcgcagctcgtgaagatcataagaagaccga 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 AACTAGGCTGCACATTCGCCCTCGAATATTTCACACTCGGAAGTATCTCGAGAGACCGA 602

QY 1338 tgccttggttatggatatagcctcttagag 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 TGTTTGGTTACGGCATATGCTCTCGAG 633

RESULT 10
A1729440 703 bp mRNA EST 11-JUN-1999
LOCUS BNLGH113383 six-day Cotton fiber Gossypium hirsutum cDNA 5' similar
DEFINITION to (U93048) somatic embryogenesis receptor-like kinase [Daucus
carota], mRNA sequence.
ACCESSION A1729440
VERSION A1729440.1 GI:5048292
KEYWORDS EST.
SOURCE upland cotton.
ORGANISM Gossypium hirsutum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
          eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 703)
AUTHORS Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.
TITLE ESTs from developing cotton fiber
JOURNAL Unpublished (1999)
COMMENT On Jun 5, 1998 this sequence version replaced gi:3188791.
        Contact: Ben Burr
        Biology Department
        Brookhaven National Laboratory
        Upton, NY 11973, USA
        Tel: 516-344-3396
        Fax: 516-344-3407
        Email: burr@bnl.1.bnl.gov
        Seq primer: 73 Primer.

FEATURES
    source          Location/Qualifiers
                        1..703
                        /organism="Gossypium hirsutum"
                        /cultivar="Acacia Maxxa"
                        /db_xref="taxon:3635"
                        /clone_id="Six-day Cotton fiber"
                        /tissue_type="Immature fiber"
                        /dev_stage="Six days post anthesis"
                        /lab_host="XIL-Blue"
                        /note="Vector: pBluescript II KS+"

BASE COUNT      188 a      148 c      180 g      186 t      1 others
ORIGIN
Query Match          13.8%; Score 250.4; DB 51; Length 703;
Best Local Similarity 65.1%; Pred. No. 5e-49;
Matches 368; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

1070 ttgattggccaactaagaagagatgcacatagatctcgtcagggtcttcttattgc 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 TAGATTGGCCTCAAAAAGAAAGATTAACCTAGTGTCAGCAGTGGCTCGGATACCTTC 68

1130 atgacatctgtatccccaagattatccatcgtgattgaagaagctgcaaatatatttgg 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 ATGAACATTGCAATCCTAAGATCATTCATCGAGATGTGAAGCTGCAAAATGATTAATGG 128

1190 acgaagaatttggaggtctgttgaagtgtatttgggttgctgtagctcatgattcaagg 1249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 ATGAAGATTTTGAAGCAGTGTGCGGTGACTTGGCTTCGGAATTGGATGTGAGGC 188

1250 ataccatgttaacatcgtctgaagggtgaaccttgggtcacaatagctcccgatcctt 1309

```



```

Db 189 GGAAGATGTGACAACTCAAGTTCGCGGACAAATGGACACATAGCACCGAGACTTGT 248
      || ||| ||||| || ||| ||| ||||| ||||| |||
Qy 1310 cgaatgaaagtcacgcagagagccgcatgtcttggtatggtatctcttagagc 1369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 CCACTGGGAATCTACTCAGAAAGAAAGAGATGCTTGGTTTAAAGGATATCTCTCAAGC 308
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1370 tcatctatgcagagagccttgatctgtcgccttcgcaacgatgatgtatgt 1429
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 TTGAGACAGGCAACGTCATTTATTTTCACCCCTGGAAAGATGAAGATGATTTCTGC 368
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1430 tcttgatggtgttaaaagccttttgaaagaaagtgtgagatgctgtgcattc 1489
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 TTTCTGACTATGTCANNAACCTGGAAGGAGAAACCTGATGCTATGCTATGATCATA 428
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1490 accctggagaaacattactctgacacagagtgagcagctattccaagtgcattactc 1549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 429 ATCTTAATTAATAATTACACATGAGAGCGTTAAAGCATATCATCAAGTTCATTTGATT 488
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1550 gtaccaggttcgcgaatgagcgccttaagatgctcagagtagtccgaatgtgaa 1609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 GCACCCCGACCTCCCAAAACCTCCGGGAATTTTGAAGTGTAAAGATCTCTGAG 548
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1610 gtgatgaccttcgcaaaagtctgga 1634
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 GAAACAGGCTCTCCCAAGGTGA 573
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
LOCUS A1486453 471 bp mRNA EST 29-JUN-1999
DEFINITION EST244774 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
cLED862, mRNA sequence.
ACCESSION A1486453
VERSION A1486453.1 GI:4381824
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 471)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.J., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Roming,C.M., Craven,M.B.,
Fuji,C.Y., Bowman,C.L., Niemman,W., Fraser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
On May 7, 1998 this sequence version replaced gi:3121325.
CONTACT: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.

FEATURES
SOURCE
location/Qualifiers
1..471
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLED862"
/tissue-type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST library. OligoDT-primed and
directionally cloned cDNA in vector lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."

BASE COUNT 136 a 88 c 112 g 135 t

```

```

ORIGIN
Query Match 12.3%; Score 223.2; DB 47; Length 471;
Best Local Similarity 67.8%; Pred. No. 1.3e-42;
Matches 312; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy 992 ggcctcttatacatcatcagcagtaagtgagtgctgcgtgtttaaagagcgc 1051
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GCCTGTACTATATACCATATATGCAAGATTTAAGTGTTGCTATGCTAGTAACTA 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1052 agccatcagaacctccctcttgattggccaactagaagaagattgacatgactgcta 1111
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AACCTGGGAGAGCGCTGTTTGAATTTGGCAATGAGAACGCTGTGGCGGTGGTACTGCAC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1112 gggggcttctcattatgacagcagcattgtgacccaagaattatccatgtgtaaaag 1171
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GTGGACTTAATTAACCTACTACAGAACCTGTAAATCAAAAGATTATTCACCGTGAATTAAG 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1172 ctgcaatatattatttgacgagaattgagcgtgttgagtgatttggttagcta 1231
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CACTATATGTGTACTAGATGAAGATTTTGAAGCTGTAGTGTGATTTGGCCTGGCAA 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1232 ggcctatgattacaagagatcaccatgctgtgttaagggtaccttgggtctaca 1291
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGCTAGTTGACGTTAAGAAACCAATGTGACTACTCAAGTTGCTGTAATGGGCCATA 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1292 tagctccgaagtaacctctgactgagaagtcacagagacgagctgtctgttatg 1351
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TAGCTCTGAAATACCTTATCCACAGCGCAATATCAAGAAATCAATGATTTTGGCTATG 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1352 ggaattatgctcttagagccattactgacagagagcctttgattctgcgccttgca 1411
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GATCATGTGCTTTTGGAAATTTGAACCGGCAACGTCATATACCTTCACGCGCTTAAGAG 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1412 acgatgatgatgttatgttctgttgattgggttaaaagcct 1451
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ATGAAGATGATGTCTTGTGCTGACCATGATCAGAAACT 460
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
LOCUS A1165643 380 bp mRNA EST 03-DEC-1998
DEFINITION A087P60u Hybrid aspen plasmid library Populus tremula x Populus
tremuloides cDNA 5', mRNA sequence.
ACCESSION A1165643
VERSION A1165643.1 GI:3856928
KEYWORDS EST.
ORGANISM Populus tremula x Populus tremuloides.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 380)
AUTHORS Sterky,F., Regan,S., Karlsson,J., Hertzberg,M., Rohde,A.,
Holmberg,A., Amin,B., Bhalerao,R., Larsson,M., Villarroel,R., Van
Montagu,M., Sandberg,G., Olsson,O., Teeri,T.T., Boerjan,W.,
Gustafsson,P., Uhlen,M., Sundberg,B. and Lundberg,J.
Gene discovery in the wood-forming tissues of poplar: Analysis of
5,692 expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)
On Jan 17, 1998 this sequence version replaced gi:2044473.
CONTACT: Sterky F
Department of Biotechnology
Royal Institute of Technology (KTH)
Teknikringen 34, S-100 44 STOCKHOLM, Sweden
Tel: +46 8 790 8287
Fax: +46 8 24 54 52
Email: fredrik@biochem.kth.se
PCR primers
FORWARD: AAAGGGGATGTGCTGCAAGCGG
BACKWARD: GCTTCGGCTGTATGTTGTG

```


st@watson.wustl.edu
is available through: Genome Systems, Inc. 4633 World
Chalice St. Louis, Missouri 63134. For further information
000) 430-0030 or (314) 427-3222 FAX: (688) 919-3324 or (314)
or contact: clones@genomesystems.com or
omesystems.com web site: www.genomesystems.com
reversed clone: similarity on wrong strand
er: 17 BT from Amerham.
Location/Qualifiers
1. 474
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1005-165"
/clone_1db="Gm-c1005"
/lab_host="X110-Gold"
/note="Vector: pBluescript II XE; Site: 1; EcoRI; Site_2:
XhoI; This library was constructed by Dr. Randy C.
Shoemaker and Dr. John Expelling, USDA-ARS Agronomy
Department, G401 Agronomy Hall, Iowa State University,

Fax: 864 656 4293

Email: dfrisch@CLEMSON.EDU

FEATURES 5 prime sequence.
Location/Qualifiers
1. .524
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEC3318"
/clone_id="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: GibcoBRL laboratory; circ - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 159 a 91 c 135 g 139 t
ORIGIN

Query Match 10.8%; Score 196.6; DB 64; Length 524;
Best Local Similarity 66.3%; Pred. No. 2.5e-36;
Matches 299; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

```
OY 1203 ggctgtttagtgatttggttagctagctcatgtatcacagaatacccatgtac 1262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 ggcgtcctggagacatttgatggcnaacatttgatcattccgattccatgtcac 60
OY 1263 aactgtctaaagggtaaccttggtacatagctcccgatcctcgaactggaagtc 1322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 AACTGCAGTTAGGGGACAGTACATATAGCTCTGATCTCTCTACAGACAGTC 120
OY 1323 atcagaagaagaccgtgtcttggttatggatattgctcttaagctcattactgaca 1382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CTCTGATTAACACGATGTTTCGGGTTTGAATTCCTCTGAGAGTTGATATCTGGCCA 180
OY 1383 gagaagcttctgactgtcgccttcgcaacgatgatgttatgttggattggt 1442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 161 GAGAGCTCTTGATTTGGTTAAGCAGCAATCAGAAAGTGTTA---TGCTGATGGGT 237
OY 1443 taagaagcctttgaaagaagaagaagttggagatcgtggtcgatccctgaacctgagaacaa 1502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 238 GAAGAGATTCACCTGGAGAAAAGCTTGATTGCTGTGAGCAAGACTTAAAGACAA 297
OY 1503 ttacattgacacagaagttgagcaagcttattcaagtaagcattactctgtaccagggttc 1562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 298 ATATGATCGGATCGAGCTTGAGGAATGCTACAGTACCTTAATATGCACTCAATCA 357
OY 1563 gccaatggagcggcctaagatgtcagaagtgatccgaatgttgaagtgatggccttgc 1622
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 358 CCTAGTCATAGACCAAAATGTGAGAAGTGTAAGATGCTCGAAGGAGACGACTTGC 417
OY 1623 agaaaagtgggcgaagtggcgaaaagttagaa 1653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 418 AGAGAAATGGGAGCTTCACAAAGACAGAA 448
```

Search completed: June 23, 2000, 19:06:03
Job time: 27396 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 23:41:42 ; Search time 12463.9 Seconds
(without alignments)
-318.517 Million cell updates/sec

Title: US-09-180-798-20
Perfect score: 4081
Sequence: 1 tctagaacctttgcatcat.....tacgaatcatgcttctaga 4081

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenBank:*

1: gb_dal:*

2: gb_dal:*

3: gb_dal:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ov:*

8: gb_ov:*

9: gb_ov:*

10: gb_ov:*

11: gb_ov:*

12: gb_ov:*

13: gb_ov:*

14: gb_ov:*

15: gb_ov:*

16: gb_ov:*

17: gb_ov:*

18: gb_ov:*

19: gb_ov:*

20: gb_ov:*

21: gb_ov:*

22: gb_ov:*

23: gb_ov:*

24: gb_ov:*

25: gb_ov:*

26: gb_ov:*

27: gb_ov:*

28: gb_ov:*

29: gb_ov:*

30: gb_ov:*

31: gb_ov:*

32: gb_ov:*

33: gb_ov:*

34: gb_ov:*

35: gb_ov:*

36: gb_ov:*

37: gb_ov:*

38: gb_ov:*

39: gb_ov:*

40: gb_ov:*

41: gb_ov:*

42: gb_ov:*

43: gb_ov:*

44: gb_ov:*

45: gb_ov:*

46: gb_ov:*

47: gb_ov:*

48: gb_ov:*

49: gb_ov:*

50: gb_ov:*

51: gb_ov:*

52: gb_ov:*

53: gb_ov:*

54: gb_ov:*

55: gb_ov:*

56: gb_ov:*

57: gb_ov:*

58: gb_ov:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	4081	100.0	4081	5	A67815	A67815	Sequence 20
2	4039.2	99.0	98471	8	AC012654	AC012654	Arabidops
3	540.8	13.3	2089	5	A67827	A67827	Sequence 32
4	470.2	11.5	88401	8	F23M19	AC007454	Arabidops
5	393.8	9.6	96475	7	ATF17M5	AL035678	Arabidops
6	387.4	9.5	6695	5	A67796	A67796	Sequence 1
7	375.6	9.2	65899	50	AC006436	AC006436	Arabidops
8	289.6	7.1	120787	8	AC002292	AC002292	Genomic s
9	277.8	6.8	86722	7	AB012245	AB012245	Arabidops
10	273.2	6.7	86064	7	AB013395	AB013395	Arabidops
11	270.8	6.6	1755	8	DCU93048	U93048	Daucus caro
12	263.2	6.4	1815	5	A67797	A67797	Sequence 2
13	262.2	6.4	84896	7	AB025639	AB025639	Arabidops
14	252.4	6.2	110211	50	AC005170	AC005170	Arabidops
15	244.6	6.0	89779	7	AB005234	AB005234	Arabidops
16	225.2	5.5	134784	8	F17123	AF160182	Arabidops
17	148.8	3.6	2278	8	SBRLK1	Y14600	Sorghum bic
18	122.2	3.0	93257	8	ATF6B4	AL034567	Arabidops
19	112.2	3.0	93353	7	ATF10M6	AL021811	Arabidops
20	112.8	2.8	42356	7	AB018120	AB018120	Arabidops
21	112	2.7	2010	8	AF024650	AF024650	Arabidops
22	112	2.7	5438	7	ATZ84202	Z84202	A. thaliana
23	112	2.7	59376	50	AC006072	AC006072	Arabidops
24	110.4	2.7	80770	7	AB005240	AB005240	Arabidops
25	108	2.6	3176	5	E12705	E12705	Arabidops
26	108	2.6	3176	8	ATU47029	U47029	Arabidops
27	107.2	2.6	94369	8	ATAC009755	AC009755	Arabidops
28	107.2	2.6	103960	8	ATAC011664	AC011664	Arabidops
29	106.2	2.6	9295	5	E12706	E12706	Arabidops
30	106.2	2.6	9295	7	ATHER	D83257	Arabidops
31	106.2	2.6	77040	50	AC004484	AC004484	Arabidops
32	104.6	2.6	75050	8	AC002334	AC002334	Arabidops
33	103.8	2.5	89338	50	AC006068	AC006068	Arabidops
34	101	2.5	78145	7	AB009053	AB009053	Arabidops
35	100.4	2.5	81736	7	AB019235	AB019235	Arabidops
36	99.2	2.4	14313	7	AP000559	AP000559	Oryza sat
37	99.2	2.4	135634	7	AP000391	AP000391	Oryza sat
38	97.4	2.4	98259	8	AC002392	AC002392	Arabidops
39	97.2	2.4	78378	8	AC007260	AC007260	Arabidops
40	97	2.4	80402	7	AB025636	AB025636	Arabidops
41	97	2.4	113567	8	AF058826	AF058826	Arabidops
42	96.8	2.4	68870	50	AC005312	AC005312	Arabidops
43	96.2	2.4	4604	7	LELRGENE	X95269	L. esculentu
44	95.8	2.3	2808	7	RICPRKI	L27821	Oryza sativ
45	93.6	2.3	89029	50	AC006569	AC006569	Arabidops

ALIGNMENTS

	RESULT	1	A67815	4081 bp	DNA	PAT	05-MAY-1999
	LOCUS	A67815	Sequence 20 from Patent WO9743427.				
	DEFINITION	A67815					
	ACCESSION	A67815					
	VERSION	A67815.1	GI:4756638				
	KEYWORDS	.					
	SOURCE	thalae cress.					
	ORGANISM	Arabisopsis thaliana					
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi-					
	REFERENCE	1 (bases 1 to 4081)					
	AUTHORS	De V.S., Schmidt E.D., Van H.G. and Hecht,V.F.					
	TITLE	PRODUCTION OF APOMICTIC SEED					
	JOURNAL	Patent: WO 9743427-A 20-NOV-1997;					
	CIBA GEIGY AG (CH)						
	FEATURES	Location/Qualifiers					
	source	1..4081					
		/organism="Arabidopsis thaliana"					
		/db_xref="taxon:3702"					
		/clone="ARABIDOPSIS SERK GENE"					
	exon	1280..1367					
	exon	1796..1928					
	exon	2014..2085					
	exon	2203..2346					
	exon	2450..2521					
	exon	2617..2688					
	exon	2772..2884					
	exon	3015..3146					
	exon	3305..3646					
	exon	3760..4081					
	BASE COUNT	1120 a 770 c 785 g 1406 t					
	ORIGIN						
	Query Match	100.0%; Score 4081; DB 5; Length 4081;					
	Best Local Similarity	100.0%; Pred.No.0;					
	Matches 4081; Conservative	0; Mismatches 0; Indels 0; Gaps 0					
Oy	1	tctgaaccccttggatcatacatgaaataagaagtcacaccacaactgggtgaacat	60				
Dd	1	TCCTGAAACCCTTTGGATCAATAAGAAATAAAGAAGCCATCACCACAATGGGTGAACAT	60				
Oy	61	aattggatatattaagggttaaccaatitaatcgcttttatctttacctct	120				
Dd	61	AATTGGATATTATTAAAGGGTAAACAATGAATCGCTTTATTATTACTTCT	120				
Oy	121	actcaaatctgatcygcagtctttttttttttaaatgataagaacaagtaactgttaa	180				
Dd	121	ACTCAAATTGANTGGCAGTTTTTTTTTTTAAGATTAAGAACAGATCTGTTTAA	180				
Oy	181	tgttgatgtgtagaagaagtagaagcataatcgycgcgccataactctccaagty	240				
Dd	181	TGTGATGTGTAGAAAGTAGTAAGATCATATCGGCACGCCATATCACTTCCACAGTG	240				
Oy	241	gaacttggccaatttgtcttcgccgctctcacagtttcttcaccaaatltttgtg	300				
Dd	241	GAACTTGCCCAATTTTGTCTTTGCCGCTCTACAGTTTCTTCCACCAATTTTGTGG	300				
Oy	301	acaaaactcaaatctttccaatctccacactctcgcacaaagtgggtttaagaagaatatacgc	360				
Dd	301	AAAAACTCAAAATCTTYCAATCTCATCTCGCCAAGTTGGGTTAGAAAAGAAATATAGC	360				
Oy	361	aaacacaaataatcttattgttcgactggtttaaccaacaacaaattcacaacattgaa	420				
Dd	361	AAACACAAATATCTTATTGTTCGCAAGGTTTATCAATCACAAATATCCAAACATTGTAA	420				
Oy	421	aaaaaattcacattttgtgtatgagatggtcccacatgatagtagaacctcttaaacatt	480				
Dd	421	AAAAAATTCACATTTTGTGTATGAATATGCCACATGATATGAACCTCTTTAAACTTT	480				

QY	481	taacttactcttcaataacggaatbaacgaatcttacttgcaataaaattgaagaag	540
Db	481	TAACCTTACTTTCATAATACGGGATTTACAACTTACTTGCAATTTAAATTTAAGAAAG	540
QY	541	gttttctacttaagaagaaagagaccacaagagagaggtttgaaccaagagaacgg	600
Db	541	GTTTTCTACTTAAAGAAAAAGGAGACCCACACAGAGAGGTTTGACAGAGAAACGG	600
QY	601	tgcatagccttaagaagcttcaactcttaactcttaccacaacccaagacgalytcaacttcaa	660
Db	601	TGCATAGGCTTAAAGAGCTTCAACTACTTATCCCAACCAAGGAGATGCATCTTTCAA	660
QY	661	ccatctctctctcccccgaacccggttttttgaacgagctcaatctcgagagaacagct	720
Db	661	CCATCTCTTCTCTCCCCGAGACCCGTTTTTTTACCGGTACGTTGCGGACGACACCGCT	720
QY	721	taaggagcactatatactctctgtctctctctctacacactgcalyccataaataag	780
Db	721	TACGGGACGCTTAAATTTCTCGTCTCTCTCTACACACTGCATGCCATTAATTAAG	780
QY	781	cccgcttagactcttaaaatatataatatatacagaagaaagctttttattcata	840
Db	781	CCCGTTGAGACTTTAAATATTAAATAATAATACAGAAACCTATTTATTTCANA	840
QY	841	agaaagaaagagagagacaacaacaacaactacataagtttctctgacagagcttgt	900
Db	841	AGAAAGAAAGAGAGAAACAACAACAACATATCATAGTTTCTCGGACGCTGTT	900
QY	901	gttggagcttaataaagaagccttctgttctatctactcaagtagaattttcccaaaa	960
Db	901	GTTGGAGCTTAATAAAGCTCTTTTGTATTCTTCCACGTAGATTTTCCCAANA	960
QY	961	gctcttatttttctgtttaaaaaaaagttcatcttattcaactttgtttacagt	1020
Db	961	GCTCTATTTTTGTTTAAAAAAAAGTTCATCTTATTCACCTTTGTGTTTACAGT	1020
QY	1021	gtgtgtgabaagagagagtggttgatttgagagaagacgagacgagacgagag	1080
Db	1021	GTGTGTGAGAGAGAGAGTGtGTGTGATTGAGAGAAAGCAGCAGAGAACGCCGGAG	1080
QY	1081	aattggatttcttatttcttactcttctgttctttaaagtaagtggtttcttaa	1140
Db	1081	AATTGGATTTTATTTTATTTTACTCTTGTGTGTTTATGCTAATGGGTTTTTAA	1140
QY	1141	aagggtttcgaaaaaatgtagtggttggttgagagtgctctgtbaagaggttaag	1200
Db	1141	AAGGGTTATCGAAAAAATGAGTGAGTGTGTGAGAGTGTCTCTTAAAGGTTTAAAG	1200
QY	1201	tgtgtatttcggaagttagaggtttcttcggaactgaagagatcaaatcaagaatcgaaa	1260
Db	1201	TGTGTATTTCGGAAGTTAGGGTTTCTCGGATCTGAAAGATCAATCAAGATTGAAA	1260
QY	1261	tttagcatgttgtttgaagaagagagtcgaattatggtgttctatcttactcaatgat	1320
Db	1261	TTTACCATTTGTTGTTGAAAGAGTGAGTATAGTGCTTTATATCTTTCACCTGAT	1320
QY	1321	cttaactccgaatcaatcaactgtagctgtctgtctgtcaatttggaaaggttcggttaac	1380
Db	1321	CTTACTTCCGAATCATTCACGTGCGTGTGTCMGCTAATTTGGAAGGTTGCGGTACT	1380
QY	1381	caattaccagaacttactcgtttctcaaatacttctcgatctcttttattcttgagagt	1440
Db	1381	CAATTACACACTTATACGCTTCTCTCAATTAATCTTCGATTTCTTTTAAATTTGGAGGTG	1440
QY	1441	aatgcgtaactttagtgtcgtcatcttgatttgaataatgttgttcttcttcttatt	1500
Db	1441	AATGCCATCTTTAGTGTGCTGCATTTTGATTATTAAGAAATGTGTGTTCTTTGTATTT	1500
QY	1501	gtaaagatttagtgcagtagtacttgaatacacgttttgccttctctgttcagatcaact	1560
Db	1501	GTAAGATTTAAGTGCTAGTACTTTGAATTAACGTGTTTGTCTTTCTTTGTCAGATCAACT	1560

QY 1561 ttgatattgaaagcagctgtcttgggttgaaaaagctgggtattgatalcttaaga 1620
|||||
Db 1561 TTGATATTGTAAGAGCATGTTCTTGGGTGAAGAGCTGGGTATTTGATATCTAAGA 1620
QY 1621 ttgatgttgatcccaaatctctcgaagaactcttggttttgggtttgtaaga 1660
|||||
Db 1621 TTGAGTGTGTGATCCAAACATTCCTGAAAGACTTCATTTGGTTTGGTTTAAAGA 1660
QY 1681 attgtttaattatagctcctaatcctcagagagagccgtttgtaagtctcctcga 1740
|||||
Db 1681 ATTGTATTATTTATTTAGCCTTAATCTCAGAGAGGCCGTTTGATTAAGTCTCTCTTGA 1740
QY 1741 attagactttcccaattgaltgtaattgtgtagattgttgttctgttatagtgat 1800
|||||
Db 1741 ATTAGACTTTCCAAATTGATGTAATTGTGTAGATTGTGTCTTGTATAGGTGAT 1800
QY 1801 gcttgacactttgaggggttaacttagttagtcaaaaagtctctgcagagctggat 1860
|||||
Db 1801 GCTTTGCACTTTGAGGGTTACTCTAGTTGATCCAAACAAATGCTTGCAGAGCTGGGAT 1860
QY 1861 cctacgcaatgaaactcttgcaatgctgctcagatgctcacttgcaaaaagcagagctg 1920
|||||
Db 1861 CCTAGCCTTAGTAATCCTTGGCAATGTTCCATGCTCACTTGCAACAGAGAAAGATGTC 1920
QY 1921 ataaagatgtaaaagctctctctactaaaccacttttaactttgacctcagcgtggt 1980
|||||
Db 1921 ATAAAGATGTAAGGCTTCTCTACTAATCCCACTTTTAAACTTTGACCTGACGCGGT 1980
QY 1981 taacgaacttttctgtctcttttgtaaaatagatgattggggaaatgcagagctatctg 2040
|||||
Db 1981 TACCAACATTTTGTGTTCTTTGTGTAATACAGTATTTGGGGAATGCAAGATTATCTGG 2040
QY 2041 ccatttaattccagagctgtgtgtcgaagaatttgcagtattgttaagtccaactat 2100
|||||
Db 2041 CCATTAGTTCAGAGCTGTGTGTCTCAAGAAATTTGCAAGTATTGTAAAGTCCACTTAT 2100
QY 2101 gcatcagctcttaacaaaacaatccaaagatttgacagaagaagcactggaactcctt 2160
|||||
Db 2101 GCATCATCTTTTAAACAAACAAATCCAAAGATTGTGACAAAGAGACGAGTTACCTTT 2160
QY 2161 tgaattgaatccttttaacaagttcttatcttaacaggagagctttacagtaaaa 2220
|||||
Db 2161 TGTAAATGAAATCTTTTAAACAAGTTCTTATTTCTTACAGGGAGCTTTCAGTAACAA 2220
QY 2221 cataactgcccagatctcctagtaactcttgaaacttgacaactagtgagttggatct 2280
|||||
Db 2221 CATTACTGGCCGATTCCTAGTAATCTTGGAATCTGCAAACTTAAAGTATTTGGATCT 2280
QY 2281 ttaactaaacagctctcgcggtcctactcgcgaatcatttgggaaagctttcaaaagctgag 2340
|||||
Db 2281 TTACTTAAACAGCTTCTCCGCTCTATTCGGGAATCATTTGGGAAGCTTTCAAAAGCTGAG 2340
QY 2341 attcttgtagagatacatatgctttaaccggctcagttacagttcttggtttaacttagt 2400
|||||
Db 2341 ATTTCGTGAGTAACTATATGCTTTACCGGCTCACTACAGCTTATGCTTTAATCTTAAGT 2400
QY 2401 ttgttccaattttgacactcttgctgtaaaaatttacaatgcaagaatgcgagcttaaca 2460
|||||
Db 2401 TTTGTTCCAAATTTTGAAGCTTTGCTGGAATAATTACATGCAAGAAATGCGGCTTAACA 2460
QY 2461 acaaacagctcagctgggtcaatctcctatgctcagacaaataatacttaacccctcaagtgt 2520
|||||
Db 2461 ACAAACAGTCTCACTGGGTCAATTTCTATGTCACCTACCAATATTTACTCCCTTCAAGGT 2520
QY 2521 tctgagctcctcatttaacttctattatgcttacttcatctcctcctaagttgattt 2580
|||||
Db 2521 TGTGAGTCTCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2580
QY 2581 gagttaatgcaacttaacttgaatgtagcaaacacagagagatcatcaaaataacagactctc 2640
|||||
Db 2581 GAGTTAATGCACTTAACCTTGATGATGCAACACAGAGATCTAATAAATAACAGAGCTCTC 2640
QY 2641 tggttcagttcctcgaaatgagctcctctcactcttcaacacccatcaggtctcatgattc 2700
|||||

Db 2641 TGGTCAAGTTCCTGCAAAAGCGCTCTCTCACTCTTCCACACCACATCAGGTTCTATGATT 2700
QY 2701 atcctcttaagtatttcaagttgtgtgtcagtgctgtaacttattctgaacttcaat 2760
|||||
Db 2701 ATCCCTTCAAGTATTTCAGTTGTGTGTCAGGTGCTCAACTTATCTGAAACCTTTTATT 2760
QY 2761 tcccttgtagctttgctataacttagacctabtggaacctgtttacaaagtcacccatgt 2820
|||||
Db 2761 TCCCTTGCAAGTTTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
QY 2821 ccttgatctcccccgtttctctccacacacctttatccaacctcccccagtttcaacc 2880
|||||
Db 2821 CCGAGATCTCCCGGCTTCTCTCCACACACTTTATTTCAACCTCCCGAGTTTCCACC 2880
QY 2881 ccgagaaagccctccctcttttagttacatttagaagaaacgagaatgaatcttctgt 2940
|||||
Db 2881 CCGAGTAAGCTCTCTCTTTTGTGTTTACATTAATTAAGAAACGAAGATGAATCTTTGCTT 2940
QY 2941 cctctgcaatccttttctcataactcaatcttgcaataaggaataaccatgactc 3000
|||||
Db 2941 CTCTGCAATCTCTTTTCTCATATAATCATCTTGCATAAAGCAATACCAATATGATC 3000
QY 3001 taattgattcaagtgagtgatagataacttggagcaatagctggtgagtgctgcag 3060
|||||
Db 3001 TAATTGATTTCAGGCGGATGATGATTAATTAATGAGCAATAGCTGATGAGTGTGTCAG 3060
QY 3061 tgtctgtctcctcttgcctgcctcctgcgaatgaccttctgtgtgtgagcgaagaagacc 3120
|||||
Db 3061 TGTCTGTTGCTCTTGTGCTGCTGCTGCAATTAAGCTTCTTGTGTGCGACGAAGAAACC 3120
QY 3121 actagatatttctctcgaatgctcctggtgagttatttctgcattagttctgtctta 3180
|||||
Db 3121 ACTAGATATTTCCTTGATGATGCTGCTGATGATTTATTTGCAATTAATTTCTGTCTTA 3180
QY 3181 gccagcaatttgttttcagagaagaatltggaacaactgttaatgaataatcaatata 3240
|||||
Db 3181 GCCAGCAATTTTGTGTTGCAAGAAAGTATGGAACAACTGTAATGAAGAAATCAATACATA 3240
QY 3241 agtcatgtttttttaaagtaacaaactctttgagtaaaatcctcgtttgcaaatctcat 3300
|||||
Db 3241 AGTCATGTTTAAAGTATTAAGTAAACCTCTTTGAGTAATTCGATTCGAAATCTCTAT 3300
QY 3301 gcaagcgaagaagatccaaagaattcaatctggaacagctcaaaagttcttcttgcgag 3360
|||||
Db 3301 GCAGCGAGAGAAATCCAAAGTTCATCTGGGACAGCTCAAGAGTTCCTTGGGGAG 3360
QY 3361 ctacaagtgcgaatgagtggttttagtaacaagaacatttggcgagaagtgtgttgg 3420
|||||
Db 3361 CTACAAGTGGCGAGTATGCTTTAGTAAACAAGAACTTTTGGCAGAGTGGGTTGGG 3420
QY 3421 aaagtctaaaggaagcgtcttgcaagcaggaactcttgtctgtctaaagagactgaaaga 3480
|||||
Db 3421 AAAGTCTAACAGGAGAGCTTGGCAGACGGAACTCTTGTGCTGTAAGAGACTGAAGAA 3480
QY 3481 gaagcaactccagtgtagaagctccagtttccaaagaagaatgagatgataagtgagca 3540
|||||
Db 3481 GAGCGAAGTCCAGGTGGAAGCTCCAGTTTCAACGAAGTGAATGATTAAGTATGGCA 3540
QY 3541 gttcctcgaaacctgttgagatgaaggttctcgtatgaacacgacgagagatgtgt 3600
|||||
Db 3541 GTTCATCGAAACCTGTGAGATTACGAGGTTTCTGTATACACGACGAGAGATTTGCTT 3600
QY 3601 gttatccttaacatgacaaatggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3660
|||||
Db 3601 GTATATCTTACATGCGCAATGGAAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3660
QY 3661 attaaacatcttgctcctctcctcaattacttgacgtgaaagtgtttttcagtcttcc 3720
|||||
Db 3661 ATTAACATCTTGTGCTCTCTCATTAATTAATTTGAGTGAAGTGTATTTTCAATGTTTCC 3720
QY 3721 ttataggttcataaatgttgtttacaaatgaacacagagagccacccgttcaaacctc 3780
|||||

Db	3721	TTTATGGGTCATATATTGTTGGTTTACACTAATGACACAGAGAGCCOACCGTCAACACTC	3780
Qy	3781	cgcttgattggcccaacgcggaagaagatcgcgctgaagctcgaagtggtgtcttacc	3840
Db	3781	CGCTTGATTGGCCAAACCGGGAAGAGATCGCGCTAGCTCAGCTCGAGTTTGTCTTACC	3840
Qy	3841	tacttgatcaactggtgataccgaaatatcatcttaccggtgactaaagacagcaaacatctctc	3900
Db	3841	TTACTGTATCACTCGGATCCGAATAATCTTCACCGGTGACTTAAGCACCAACATCTCT	3900
Qy	3901	tagacgaagaattctgaagcggtgtcttgagaatttcggttcggttcgcaagcctaatsgactata	3960
Db	3901	TAGACGAAGAATTTCGAAGCGGTTGTTGGAGATTTCGGTTGGCAAAAGTAAATGACATTA	3960
Qy	3961	aagacactcacgtgacaacagcagctccgtggaacatcggttcaactcgctccagaatc	4020
Db	3961	AAGACACTCACGTGACAAACAGCAGTCCGTGGCCACATCCGTCACTCTCTCAGATATAC	4020
Qy	4021	tctaacccggaatatcttcagagaacacgacgtttcttgatagaagaatctgctcttag	4080
Db	4021	TCTCAACCGGAATACTTTCAGAGAAAACGACGCTTTTCGGATACGGAATCATGCTTTAG	4080
Qy	4081	a 4081	
Db	4081	A 4081	
RESULT	2		
AC012654			
LOCUS	98471 bp	DNA	PLN 10-DEC-1999
DEFINITION	Arabidopsis thaliana chromosome 1 BAC F14023 sequence, complete		
ACCESSION	AC012654		
VERSION	AC012654.2	GI:6554462	
KEYWORDS	HTG.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
REFERENCE	Arabidopsis		
AUTHORS	1 (bases 1 to 98471)		
	Liu,S.-X., Lee,J.M., Sakano,H., Yu,G., Jhaveri,A., Ienz,C.,		
	Toriumi,M., Chin,C., Chlou,J., Choi,E., Gonzalez,A., Howng,B.,		
	Koo,T., Li,J., Liu,A., Pham,P., Vaysberg,M., Altati,H., Brooks,S.,		
	Buehler,E., Chao,O., Conn,L., Conway,A.B., Hansen,N.,		
	Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Nguyen,M., Palm,C.,		
	Shinn,P., Tambunga,G., Davis,R.W., Ecker,J.R., Federspiel,N.A. and		
	Theologis,A.		
TITLE	Arabidopsis thaliana chromosome 1 BAC F14023 sequence		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 98471)		
AUTHORS	Theologis,A.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (02-NOV-1999) Plant Gene Expression Center, 800 Buchanan		
AUTHORS	Street, Albany, CA 94710, USA		
TITLE	3 (bases 1 to 98471)		
JOURNAL	Theologis,A.		
	Direct Submission		
	Submitted (10-DEC-1999) Plant Gene Expression Center, 800 Buchanan		
	Street, Albany, CA 94710, USA		
COMMENT	On Dec 10, 1999 this sequence version replaced gi:6175131.		
FEATURES	Location/Qualifiers		
SOURCE	1. 98471		
	/organism="Arabidopsis thaliana"		
	/cultivar="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="1"		
	/clone="F14023"		
BASE COUNT	31501 a 17757 c 17475 g 31738 t		
ORIGIN			

Query Match	99.0%	Score 4039.2	DB 8	Length 98471
Best Local Similarity	99.7%	Pred. No. 0		
Matches 4070	Conservative	0	Mismatches 3	Indels 9
Gaps 2				
QY	1	tctaaacocctttatataatgaataaagagtcacacacagctgggtagat	60	
Db	76556	tctaaacaccttttatcatatgaataatgaatccatccacacatggggtagat	76615	
QY	61	aatgctgatatlaaaggtacaaatgtaacctgcctttatcttacttaccct	120	
Db	76616	aatgctgatatattaaaggtacaaatgtaacgctttttattttacttttaccct	76675	
QY	121	actcaaatgtatggcag-tttttttttttttaaataagaacatctgttta	179	
Db	76676	actcaaatgtatggcagttttttttttttttaaataagaacatctgttta	76735	
QY	180	atgcatgtgtagtaaacagtagtaagtcatatcggaagccatctacttccaggt	239	
Db	76736	atgcatgtgtatgaacaaagtagtaaatcatatcgccagcccatctacttccacagt	76795	
QY	240	ggaacttggccaaatctgtcttcgcgtctccacagcttcttccacaaatcttgtt	299	
Db	76796	ggaacttggccaaatctgtcttcgcgtctccacagcttcttccacaaatcttgtt	76855	
QY	300	gacaaactcaaatcttcaatctcatctctgcgaagtctgggttagaagaatatacag	359	
Db	76856	gacaaacacccaatcttcaatctcatctctgcgaagtctgggttagaagaatatacag	76915	
QY	360	caaacactatctctttatgtcatagtttatcaatcaaaaaattcaaacacttgta	419	
Db	76916	caaacactatctctttatgtcatagtttatcaatcaaaaaattcaaacacttgta	76975	
QY	420	aaaaaaatcacaattcttgatagaatggctcacatgataagtaaacctttaact	479	
Db	76976	aaaaaaatcacaattcttgatagaatggctcacatgataagtaaacctttaact	77035	
QY	480	ttaacttacttcaataatacaggaattacgaatcttctgcatlaaaatttgaanaa	539	
Db	77036	ttaacttacttcaataataacaggaattacgaatcttctgcatlaaaatttgaanaa	77095	
QY	540	ggttttctacttaagaaaaaaggaccocacagagagaggtttgaccagagaanaag	599	
Db	77096	ggttttctacttaagaaaaaaggaccocacacagagagaggtttgaccagagaanaag	77155	
QY	600	ggtcatagccttaagagcttcaactacttaccaccaaccacaagcgatgcatctta	659	
Db	77156	ggtcatagccttaagagccttcaactacttaccaccaaccacaagcgatgcatctta	77215	
QY	660	accatctctctcccccgaaccgcttttttgaacggtcagttcgggcagcagcagcg	719	
Db	77216	accatctctctcccccgaaccgcttttttgaacggtcagttcgggcagcagcagcg	77275	
QY	720	ttaagggcagcttatatctcgtcttctcctcctcaacccaactgatatgccataataa	779	
Db	77276	ttaagggcagcttatatctcgtcttctcctcctcaacccaactgatatgccataataa	77335	
QY	780	ggccgttgagatctttaaataatcaataataatacaagaaagcattttatcat	839	
Db	77336	ggccgttgagatctttaaataatcaataataatacaagaaagcattttatttcat	77395	
QY	840	aagaagaanaagagagaacaacaacaacatacatagtttctctgycagcgttgt	899	
Db	77396	aagaagaanaagagagaacaacaacaacatacatagtttctctgycagcgttgt	77455	
QY	900	tgttcgcgctataataaagcctcttggttatatacttaactcaagatatttcccaaaa	959	
Db	77456	tgttcgcgcttatataaagcctcttggttatatacttaactcaagatatttcccaaaa	77515	
QY	960	agccttatatttttggtttaaanaaaagttcaccttatacaactttgtttaag	1019	
Db	77516	agccttatatttttggtttaaanaaaagttcaccttatacaactttgtttacag	77575	
QY	1020	tgttgtgtgtagagagagtggttggttatgtagaagaacgacagcagagacgcgga	1079	

Db	77576	TGTGTGTGAGAGAGAGAGTGGTTGATTGAGGAAGAAGACGAGACGAGAAAGCCGGA	77635
Qy	1080	gaataagaattttatatttttttaactcttctgttttttaataatgctaagggtttta	1139
Db	77636	GAATTAGATTTTGAATTTATTTTACTCTTGTGTTTAAAGCTAAAGGGTTTTTA	77695
Qy	1140	aaagggtatcgaaaaaatgagtgagttgtgttgagggtgtctctcgtgaaggtttaag	1199
Db	77696	AAAGGTTATCGAAAAATGAGTGAAGTTGTGTGAGGTGTGCTCTGTAAGGTGAATG	77755
Qy	1200	gtgtgatttttcggaagttaagggtttctcgagatcgaagaagatcaaatcgaa	1299
Db	77756	GTGGGATTTTCGGAAGTGAAGGTTTCTCGAGATCTGAAGATCAAAATCAAGATTGCAA	77815
Qy	1260	atttgacattgtgtgtgaataagtgagtgagttatgtgtgtttatacttaacttaacga	1319
Db	77816	AATTAGCATGTGTTTGAAGAAGAGTGAATGTGTGTTTACTTTCACCTGA	77875
Qy	1320	tcttaactccgatatcatcaactgtgactgtctctcgaatttggaaggttcgtgttac	1379
Db	77876	TCTTACTTCGGAATCATTTACGTGTGGCTTGTGCTAATTGGAAGGTGAGGTATAC	77935
Qy	1380	tcaattactcagcttaactcgtttctcaattactcttcogactcttttatcttgagagc	1439
Db	77936	TCAATTACTCACGCTTTACTCGTTTCTCAATTTCTTCGATTTCTTTTATTTGAGAGT	77995
Qy	1440	gaatgcatactcttagtggtgtgcattttgatttaagaanaatgtgtgttctctgtat	1499
Db	77996	GAATGCGATCTTGTAGTGTCTGATTTGATTTGAAAATTTGTGTTGTTTGTGAT	78055
Qy	1500	tgtaaagatttagtgctaacttaactgaatacactgtttgtcttctcgttcgaacac	1559
Db	78056	TGTAGATTTATGTGCTATCTTGAATACACTGTTTGTGCTTCTGTTCAGATCAAC	78115
Qy	1560	tttgtatattgtaaggcagtctcttggtgtgaagaagctgggttaattgabatctaag	1619
Db	78116	TTTGTATATTGTAAAGGCGATGTTCTTTGGGTGAAGAAAGCTGGGTATTTGATATCTT	78175
Qy	1620	atttgatgtgttgatccaacaactctctgaagaacttaactgtttgtgtttgtaag	1679
Db	78176	ATTGATGTTGTGATCCAAACTTCTGTAAAGACTTCATTTGTTTGTGTTTGTAAAG	78235
Qy	1680	aatttgtaattatagaccttaactcaagagagcctgtttgaatagttactctctga	1739
Db	78236	AATTGTATTAATTATAGCTCTAATCTCAGAGAGCCGTGTAATGTTCTCTCTGA	78295
Qy	1740	aattgaactttccaacaattgataatgtgtagaattgttctgtctgtataagfga	1799
Db	78296	AATTAGACTTTTCCAAATTTGATGCTAATTTGTATTTGTGTTATTTGATGCTGA	78355
Qy	1800	tgtctgtgaactttgagggttaacttagttatccaacaagaatgtcttcagaagctgga	1859
Db	78356	TGCTTGTGCACTTGTAGGGTTACTGTAGTGTGATCCAAACAAAGTCTTCAAGCTGGGA	78415
Qy	1860	tccctaagctagtaatccctctgcacatggttcacatgcttaacttcgaacaagaagagtg	1919
Db	78416	TCTTAGCTAGTAATCTTGTGCAATGTTCCATGTCACTTTGAACAAGAGACAGTGT	78475
Qy	1920	cataagagtgtaagaacttctcttaactaatcccaacttttaaaccttgaaactcaagctg	1979
Db	78476	CATTAAGAGTGAAGCTTCTTCTACTAATCCCACTTTTAAACCTTTGACCCAGAGCTGG	78535
Qy	1980	tttaacgaacttttctctcttctgtcaaatagagattggggaatgagaagttatctg	2039
Db	78536	TTTACCGACATTTTGTGTTCTTTTGTCAAAATACAGTATTTGGGGAATGCAAGATTATCTG	78595
Qy	2040	gcaatttagttccaagagctgtgtgtctcaagaatttgcaagtaattgttaagtccaacta	2099
Db	78596	GCAATTATAGTTCCAGAGCTGTGTGTCTCAAGAAATTTGCGATTTGTAAATTCCACTTA	78655
Qy	2100	tgcatacgtctttaacaacaataccaaagattgaacagaagaagcaactggaattacct	2159

Db	78656	TGCATCATGCTTTAAACAAAACAAATCCAGATTGTGACAGAGAGACCTGAGTTACCT	78715
Qy	2160	ttgtaattgaatctctttaacaagttcttactttcttaaggagcttaacagtaaca	2219
Db	78716	TGTATTTGAAATCTTTTAAACAAGTTCTTATTTTCTTACAGGAGCTTTACACTAACA	78775
Qy	2220	acataactgcccgaattctcagtaatcttggaatctcgaaactgaactagtgagttgagc	2279
Db	78776	ACATTAAGTGGCCGATCTCAGTAATCTTGAAATCTGCAAACTAGAGTGAAGTGGATC	78835
Qy	2280	tttaactaaacagcttctccggtcccatcccggaatcaattgggaagctttaagcga	2339
Db	78836	TTTACTTAAACAGCTTCTCCGGTCTTATTCGGAATCATTTGGGAAGCTTTTAAAGCTGA	78895
Qy	2340	gattctgtgatatcatatagtttcaacggctcaagttcaactcttgtttaacttagg	2399
Db	78896	GATTTGCTGAGATACATATGCTTTACCGGCTCACTTACAGTCTTTGTTATCTTAGG	78955
Qy	2400	ttttgtccaattttgacactttgtctgaanaatttaactgaagaatagccgcttaac	2459
Db	78956	TTTGTTCCAATTTTGAAGCTTGTGCTGAATAATTACATGCAAGAAATAGCCGCTTAAC	79015
Qy	2460	aaacaagcttcaactggtgtcaatctctatgtcaactgaaccaatatacttaacttaag	2519
Db	79016	AACAACGCTCACTGGGTCAATTCCTATGTCACTGACACCAATTTATCTACCTTCAAGTG	79075
Qy	2520	tttgtgaactcctcaatcaacttcaattatgtcttaactatctccctcaagttagtg	2579
Db	79076	TGTGATCCTCTCATTTAATCTTCAATTATGATGATGATGATGATGATGATGATGAT	79135
Qy	2580	tgaattgaagcttaactgaacttgatggaacacagaagatcbataaataacagactc	2639
Db	79136	TGAATTAATGACATTAACCTTGATGATGATGATGATGATGATGATGATGATGAT	79195
Qy	2640	ctggttaagttctctgaaagagctcctctcaacttcaactcaaccatcaagttctaagt	2699
Db	79196	CTGTTGATGCTCGAATAAGGCTCTTCTCACTCTTCAACCCATCAAGGTTCTATGAT	79255
Qy	2700	tatcctcctcaagttattcaagttgtgtgtgcaagtgtcgaacttattcgaaacttcat	2759
Db	79256	TATCCTCTCAGTATTTATGATGTTGTGTGTGATGATGATGATGATGATGATGAT	79315
Qy	2760	ttcctgtgcagtttgcttaactaactagaactaagtgaactgttaacaagtaeccatg	2819
Db	79316	TTCCCTTGTGAGTTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	79375
Qy	2820	tcctgagatcccccgttttctctcacaacactttattcaaccccccagtttccag	2879
Db	79376	TCTGTGATCTCCCGGTTTCTCTCCACCACTTTTATTCAACTCCCAAGTTTCCAC	79435
Qy	2880	cccgaagtaagcctcctcttttagtttacaattaagaacaagaagataactcttgct	2939
Db	79436	CCCGAGTAAGCTCTCTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTAT	79495
Qy	2940	tctctgtcaatcccttttctcaataaactcaactcttgcaatgaagcaataccaatgat	2999
Db	79496	TCTGTGCAAACTTTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	79555
Qy	3000	ctaatttgattcaagtggtgtatgataactgagaagaatagctgtgtgaggtgtgcgag	3059
Db	79556	CTAATTTGATTTCAAGTGTGGGTATGTAACTGTGAAGCAAAAGCTGTGGAGTTGCTGCA	79615
Qy	3060	gtgctgactgtcctctctgtcctcgtcaataagccttgcgttggtgcgagagaagaagc	3119
Db	79616	GTGCTGCTTCTCTTGTGCTCTCTCTCAATAGCTTGTGTGTGTGTGTGTGTGTGTGT	79675
Qy	3120	caactagabatttctcgaatgtgcctggtgagttatattatcgaattagttctgttctt	3179
Db	79676	CACATAGATATTTCTTGAGATCTCCCTGTGAGTTATTTATTTCCATTAATGTTCTGTTCT	79735
Qy	3180	agccaagaattttgttgagaagaatatttggaacaacaggttaatgaanaataaataat	3239
Db	79736	AGCCGCAATTTGTTTGTGAGAAAGTATTTGGAACAACGTGTTAATGAATAATCAATACAT	79795

QY 3240 aagtcattgtttttaagttacaacatcttttgtagtaaaatctcgatgcaaaatctcta 3299
 |||||
 Db 79796 AAGTCATTGTTTTTAAATTACAACTCTTTTGAGTAAATCTCGATTGCAAAATCTCTT 79855
 |||||
 QY 3300 tgcagcgaagaagatccagaagttcatctcggaacgtcaagaaggtttcttcgggga 3359
 |||||
 Db 79856 TGCACCCAGAGAAATCCAGAGAGTTCACTGCGACACTCAAGAGCTTTCTTGGCGGG 79915
 |||||
 QY 3360 gctacaagttggagttgtagtttaagtaacaagaacattttggcgaggtttgg 3419
 |||||
 Db 79916 GCTAAAGATGGCGAGTGTATGGCTTATGACAAACAACTTTTGGCGAGAGTGGCTTGG 79975
 |||||
 QY 3420 gaaagttcaagaagcgcttggcagacggaactcttggctgctgtaagaagactgaagga 3479
 |||||
 Db 79976 GAAAGCTTACAAGGAGGACCTTGGCGAGAGAACTTGTCTGTCAAGAGACTGAAGGA 80035
 |||||
 QY 3480 aagaggaactccagttgtagaagctccagttccaagaagaagtagaagtagaagtaggc 3539
 |||||
 Db 80036 AAGAGGAAGCTCCAGGTGAGAGAGCTCCAGTTCAACAAAGAACTAGAGATGATAGTATGGC 80095
 |||||
 QY 3540 agtcaatcgaaacactgtttgagaattcagaagttctgtatgacacacggaacggaagattgct 3599
 |||||
 Db 80096 AGTTCATCGAAACCTGTTGAGATTACGAGGTTCTGTATGACACCGAGAGAGATTGCT 80155
 |||||
 QY 3600 tctgtatccttcaagcgacatggaagtgctgcttgcgtcctcagaaggttaaaactaac 3659
 |||||
 Db 80156 TGTGTATCTTACATGCGCAATGGAAGTGTGCTTGTCTGACAGAGTAAAGAC---- 80211
 |||||
 QY 3660 aattaacactctgtgctcctcctcaatcttgaacgtgaagtttttcatgcttttc 3719
 |||||
 Db 80211 ---TAAACATCTTGTGCTCTCTCTCAATTACTTTGACGTGAAGTGTTTTTCATGTTTTC 80267
 |||||
 QY 3720 cttatgggttcaataatggttggttaacataatgacagaagggccaacgttacaacct 3779
 |||||
 Db 80268 CTTTATGGGTTTCAATATTGTTGTTACACTAATGACACAGAGAGGCCACCTCAACACT 80327
 |||||
 QY 3780 ccgctgtgattggcgaacgaggaagaatcgcgctgaagctcgaagctcgaagtttcttacc 3839
 |||||
 Db 80328 CCGCTTGATTGGCCAAACGGGGAAGAGATCCGGCTAGGCTCAGCTCGAGGTTTCTCTTAC 80387
 |||||
 QY 3840 ctacatgacactcgatcgatccgaagaatcactcaacgtgacgttaaaagcagaacatctctc 3899
 |||||
 Db 80388 CTACAGATGACCTGCGATCCGAGAGATCTACCTGACGTAAGCAAGCAAAATATCTCT 80447
 |||||
 QY 3900 ttgagcgaagaattcgaaacggttcttgagaattcggggttggcgaagaactaatgacctat 3959
 |||||
 Db 80448 TTGAGCAGAAGATTTCGAACCGGTTTCTTGAGATTTCGGGTTGGCAAACTAATGACTAT 80507
 |||||
 QY 3960 aaagacactcaagctgacaacagcaagtcogtggacacatcggttcaacatcgctcagaatat 4019
 |||||
 Db 80508 AAAGACACTCACGTGACCAACAGCAGTCCGTGACCATCGGTCACTCCAGAAATAT 80567
 |||||
 QY 4020 ctctcaacgggaagaatcttcagagaagaacggcgcttccgatacgaagatactctcta 4079
 |||||
 Db 80568 CTCTCAACGGGAATCTTCAAGAAACCGACGTTTTCGATAGGAAATCATGCTCTTA 80627
 |||||
 QY 4080 ga 4081
 ||
 Db 80628 GA 80629

eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsis.
 1 (bases 1 to 2089)
 REFERENCE
 AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
 TITLE PRODUCTION OF APOMICRITIC SEED
 JOURNAL Patent: WO 9743427-A 20-NOV-1997;
 CIBA GEIGY AG (CH)
 FEATURES
 SOURCE
 1..2089
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="SEER GENE CDNA"
 195..2072
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CA42254.1"
 /db_xref="GI:4756650"
 /translation="MESSYVVFILSLILPNSHMLASANLEGDALHTLRVLVDPN
 NTVQSDPPLVNPCTWTFHYTCNNNSVIRVDGNALSGHVLPELGVILKNLYELVS
 NITGPIPNLGNLNTLVSLDXLSFSGPPIESIGKSLRPLNNNSLIGSIPMS
 LNTITLQVLDLSNNRLSGSVPDNCSFSLFTPISFANNLIDCGPTSHPCGSPSP
 PPPPIPPPVSTPSPGIGTGAIVAGVAAALPEAPALAFAMRRKSPDLDFDPA
 EEDPEVHLQKRFSLRELVASDESNKNIIRGSGFKVYKGRADSTLVARLKE
 ERTPGELQFOIEVEEMISMAVHRNLRGFCMPTEBLVLYPYMANGSVASCLERP
 PSOPPLDMPFKRIALGASGLSYLHDCDPIIRRDYKANIILDEFEVAVDFFGL
 AKIMDYKDEVTAVRTGHTIAPETLSGSKSEKTEVDYGVGIMLELITGQAFDLA
 RLANDDVMILDMVGLKLEKRLKEMLVDPDIOTNFERRELDVIOVALITCGSPMER
 PKMSVRYRLIEGDGLAEKMDENQKYLEIRZELIDLSPNNSDMILDTYNLNAVELSGP
 R"
 BASE COUNT 569 a 419 c 502 g 599 t
 ORIGIN
 Query Match 13.3%; Score 540.8; DB 5; Length 2089;
 Best Local Similarity 85.3%; Pred. No. 1.7e-92;
 Matches 665; Conservative 0; Mismatches 2; Indels 113; Gaps 1;
 QY 3302 cagcgaagaagatccagaagttcatctggagacagctcaagaaggtttcttcggggaac 3361
 |||||
 Db 1018 CTGCGGAAGAGATCCAGAACTCATCTGGAGACGCTCAAGAGGTTTCTTTCGGGAGC 1077
 |||||
 QY 3362 taacaagtgcgagtgagtggttttagtaacaagaacatttggcgagaagtggttgga 3421
 |||||
 Db 1078 TACAAGTGGCGAGAGATGCGTTTGTAGTAAAGAACATTTTGGCGAGAGGTTTGGGA 1137
 |||||
 QY 3422 aagttacaagaagcgcttggcagaacgaaactctgttctgctcgaagaagctgaagaag 3481
 |||||
 Db 1138 AGCTTACAAGGAGACGCTTGGCAACGGAACCTGTGTGCTGTCAGAGACTGAAGAG 1197
 |||||
 QY 3482 agcgaactccagtgagagctccagtttcaacaagaagtagagatgataagtagcgag 3541
 |||||
 Db 1198 ACGGAAGCTCCAGGTGAGAGCTCCAGTTTCAAAAGAGAGTAGATTAAGTATGGCAG 1257
 |||||
 QY 3542 ttcatcgaaacactgtttgagattacgaagtttctgtatgacaacggaacggaagattgcttg 3601
 |||||
 Db 1258 TTCAATCGAAACCTGTTGAGATTACGAGGTTTCTGTATGACACCGAGAGATGCTTG 1317
 |||||
 QY 3602 tgaatccttaactggaagtggaagtgcttgcgtgctcagaaggttaaaactaaca 3661
 |||||
 Db 1318 TGTATCTTACATGCGCAATGGAAGTGTGCTTGTGCTCAG----- 1361
 |||||
 QY 3662 ttaaacatcttgctcctcctcaattacttgcggaagtggttttcatgtttccct 3721
 |||||
 Db 1361 ----- 1361
 QY 3722 ttatggttcataatgttggtttacaactaatgacacagaagggccaacgttacaacctcc 3781
 |||||
 Db 1361 -----AGAGAGGCGAACGTCACACAACTCC 1384
 |||||
 QY 3782 gcttgatggccaagcggaagaagatcgcgctaggtcagctcgaaggttggttctacct 3841
 |||||
 Db 1385 GCTTGATTGGCCAAAGCGGAAGAGATGCGCTAGGCTCAGCTGAGGTTGTCTTACCT 1444

Qy	3703	tggttttcaatgttttcccttaaggttcataat-----tggtgttaacactaatg	3753	/codon_start=1
				/product="putative protein"
				/protein_id="CAB38783.1"
Db	72927	TTTTGATCTTATCCATTTTAAACTTTCATACATTAAGTAGAGCTTTCTTGAATG	72868	/db_xref="GI:4490292"
Qy	3754	a-caacagagggccaccgcacacccgcgcgttgattgagccaacgcggaagagacgcg	3812	/translation="MGVEIQSPQESSTVEQLVALNPENPILPDLENVYVNTSQTY
				SLFVNLCILRLXQFEPRMNTIHVAIIVKALMAPTPDFSLFLIEPRYQMEQFK
Db	72867	ATTAACAGACAGCTCCACATACAGAGTTGCCCTTGACCTGGCAATAGACAGCAATCGGG	72808	SLIVLSHYLETGFOOFMDCAKNRHILAEVAGFOAIGAVASHLSISYQKVPSTVL
				AEAVNDGASIDKFIQOYVNTSGWIVYKSGSIVLPQNFNHPDLKKNNGENVPLEHI
Qy	3813	ctagagctcaagctcgagggttgcttaccatcatgatcactcgatccgaagatcacc	3872	ARIFPLIG"
				/complement(2572..2754)
Db	72807	CTAGAGATCACGAGGGGTTGTCTTATCTTCATGATCATGACCCCAAAATTATTAC	72748	/number=1
Qy	3873	cgtacagtaaaagcagcaacacccctcttagacgaagaatgcgaacggttcttgagat	3932	/complement(2755..2837)
				/number=2
Db	72747	CGTGATGTGAAGCTGTATAATCTGTTGGACGAGGAATTTGAGCGGGGTAGTGAT	72688	/complement(2838..3031)
				/number=2
Qy	3933	ttcgggttggcaaaagctaagactataaagacactacgcgtaacacagcagtcgtggc	3992	/complement(3032..3098)
				/number=3
Db	72687	TTGGGGTAGCTAGACTTAAGACTATAAGATACATCATGCAACAGGCTGTGCTGGG	72628	/complement(3099..3179)
				/number=3
Qy	3993	accatgcgtacatcgctccagaatatcttcaacgcggaacatcttcagagaacacgcg	4052	/complement(3180..3230)
				/number=4
Db	72627	ACTATTGGACACATGTCCTGAGTATCTCTCAACTGGAAATCTCAGAGAAACTGAT	72568	/complement(3231..3460)
				/number=4
Qy	4053	gtttcggatacgcgaatcatgcctctaga	4081	/number=5
				/complement(3461..3574)
Db	72567	GTTTGTGGCAGGATCATGCTTTTGA	72539	/number=5
				/complement(3575..3667)
				/number=5
				/complement(3668..3724)
				/number=6
				/complement(3725..3814)
				/number=6
				/complement(3815..3928)
				/number=7
				4544..6491
				/gene="F17M5.20"
				/complement(4544..4687)
				/number=1
				/complement(join(4544..4687,4781..5362,5443..5712,
				5815..6093,6423..6491))
				/note="strong similarity to WD-repeat protein -Daucus
				carota, PID:g2253631; Contains Trp-Asp (WD-40) repeats
				signature (VLAIALDHTVYLMDA) [IGGGGGDRIKFMNT]
				[VASAGDETRFNVN]"
				/codon_start=1
				/product="WD-repeat protein-like protein"
				/protein_id="CAB38784.1"
				/db_xref="GI:4490293"
				/translation="MDAGLNRCPLQEHFLPRKSKENLDREIFNRSAANFDVAHFALT
				ERRGKDQASATVSSPSKEAYRKQDAETMNIHTRILAFNRKPQAPVELLPNSASALH
				QQPKSVPRRTIPQTSERTIDADVIDDFYINLDWGSANVALIALDHTVYLMQDASTG
				STSELVIDEKGFVTSINNAPOGRHVAAGLNINSEYQIMDSANROLRLTKGQHSRV
				GSVMNNHILITGMDGLINNDVRIASPIYETRYGHTQVCGLSKSGSOOLASGN
				DVAVIHMDRSVASSNSTQWHLRHEHTSAVKLANCPQANLATGGGGGDDTIETW
				NTHGACLNSTYDNGSOYCSILMSKNEELISGCFQNOCLTIKYSVMYMEIQT
				SRVIAQSPDGCTVASAGDETRFNVNFGVPEYAKKAPKAVAEPSHVRIR"
				/complement(4688..4780)
				/number=1
				/complement(4781..5362)
				/number=2
				/complement(5363..5442)
				/number=2
				/complement(5443..5712)
				/number=3
				/complement(5713..5814)
				/number=3
				/complement(5815..6093)
				/number=4
				/complement(6094..6422)
				/number=4
				/complement(6423..6491)
				/number=5
				7856..9901
				/gene="F17M5.30"

[illegible]

REFERENCE	TITLE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	RESULT
QY 3598	cttgatcatcttaacatggccaaatgaagatggtcttcgcgtccagagataaacctaa	Arabidopsis thaliana	AC002292/c	Genomic sequence of Arabidopsis BAC F8a5, complete sequence.	AC002292	1		8
Db 64324	cttgattatccctacatggcctaatggaagcttctcttctgttttaagagctaccacatttc	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 3658	acaattaaacatcttgctcctctctccataacttgaacatgacgtgaagttttttcattg	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64384	tttgctctaaagattatttgctatattttttgtttgtttgacacgtttgagattggt	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 3718	tccttatagggtcatcaattgtttgttcaactaatgacacagagagggccacgtccaaac	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64444	cagttggttaagttatctttatctttgattt---ttctattatagacgcctcgaagacaaac	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 3778	ctcgcgtgatgtggccaaagcgagagagaatcgcgcctcagctcgcgcgagttgctt	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64500	cagcccttgctcctggccaaagaaagacatttctctctgggattacgaagggcctgcgatt	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 3838	acctacatgctacgtcgcgtccgaagatcattcacgcgtgacgttaaaagcagcacaatcc	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64560	attttacagcatctttgctcgaacaaagatcattcactcgtgacgttaaaagcagcacaatcc	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 3888	tccttaagcagaagaattcgaagcggtgtttgtgaagatttcgggtgttgcgaagctaatgact	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64620	tgtttgattgtaaggtttgtaagctgtttgttgagattttggcctcgaacaaattatgaatt	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 3958	ataaagacatcgcgtgacaaacagcagctcgttgagacatcgtgacatcgcgcacgaat	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64660	attacagcatctccattgtaacactcctctgacgggagacattgacccatattgaccccgagct	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 4018	atctctcaacccgaaatcttcagaagaaacacgcagtttccgatacgaatcatgcttc	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64740	actctcgcagcagaaattcttcgaaagactgattgtttgggtacggcgatcgttc	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
QY 4078	taga 4081	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8
Db 64800	tcga 64803	Arabidopsis thaliana	AC002292.1	GI:2252639	AC002292.1	1		8

JOURNAL	Submitted (12-JUN-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	4 (bases 1 to 120787)
AUTHORS	Federpiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R., Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Oj1.O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskala,V., Yu,G., Becker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUL-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	5 (bases 1 to 120787)
AUTHORS	Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Oj1.O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskala,V., Yu,G., Becker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (11-JUL-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	6 (bases 1 to 120787)
AUTHORS	Federpiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R., Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Oj1.O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskala,V., Yu,G., Becker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (02-OCT-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	On Jul 11, 1997 this sequence version replaced g1:2251217.
FEATURES	Location/Qualifiers
source	1. 120787 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" 177..2876 /gene="F8A5.1" join(177..214,632..707,1132..1152,1283..1479,1617..1761,1845..2159,2256..2526,2788..2876) /gene="F8A5.1" /note="Hypothetical protein" /codon_start=1 /protein_id="AAB71954.1" /db_xref="GI:2462735" /translation="MCCRWTLRONDGTEVSLVFGSMQVLAPIIGFFKQIIVLQI LNTVLDAVKGSTSPGNOTQYFAVNADKPCPTASNLERKSGLEDYVLRANCLDT MCDYCFSDREHAKYSGTGOEDKHKRGKGTMEVYVLSLSTEOHCRSGKREY LYRNPSPSPYPHALSAALKKIDPKKAGLILAVNDDGHVLEMPNFSTYQIQIAL HWINQPTKONGPGISLAKGKKNLNLKAKHGFLLSHSRKICSYVDLANSI AGLIFFSSTDIDFGDCSLVGFQIQVEERTYVNTQKTIYVIGNERKPKQDEAA PFLFDESETSFEDEVEDRENSHEDATKATKATYOFKCIATMIQE"
gene	complement(3171..4444) /gene="F8A5.2" complement(join(3171..3389,3471..3605,3706..4029,4418..4444)) /gene="F8A5.2" /note="nearly identical to rice water stress induced protein gp1d26537 537404"
CDS	protein gp1d26537 537404 /codon_start=1 /db_xref="GI:2462751" /protein_id="AAB71970.1" /translation="MAPEISVNPYITSEKHAQAPPRRAYTPIAGNDYKGVGLAK GLRWKSAIVLVAMLPDVEHEHREILRSQCVAREIEVPPDNOVERAMAYIVNI SKRIWNEEYSKMIYLDADIQVNDLHFDLSDAFYVMDCFCKEYSHSLQISI GYCOCEKVPWPEDESPPLPLTFNAGMVFSPSPLYESLQTLQLEITPPSPFADQ FLNMFPERVYKPIPLVYNLVLAAMLRBENVELEKVVHVYCAAGSKPMRYGEENM DREDIKMLYKMDVYNDESIDFSKIPADAEEYVYKSSILASVLEPMTYFAPASAA"
gene	complement(8347..9066) /gene="F8A5.3" complement(join(8347..8443,8539..8657,8941..9066)) /gene="F8A5.3"
CDS	

gene	/codon_start=1 /product="putative ADP-ribosylation factor" /protein_id="AAB71955.1" /db_xref="GI:2462736" /translation="MGISPAKLEISRLFAKKERILMVGLDAAGKTTILYKLGIIDE LRAVLIYFANKQDLPNANNAEITDKLGLSHLRQHRHYIQITCATTEGILEGMDL SNNIANKVSHF"
CDS	complement(9782..14180) /gene="F8A5.4" /note="Atyps34" complement(join(9782..9919,10031..10178,10275..10418,10509..10663,10780..10917,10991..11164,11286..11477,11815..11936,12047..12182,12300..12395,12475..12622,12803..12981,13104..13313,13414..13572,13659..13755,13867..13963,14069..14180)) /gene="F8A5.4" /note="location of EST gb w43464" /codon_start=1 /product="phosphatidylinositol 3-kinase" /protein_id="AAB71971.1" /db_xref="GI:2462752" /translation="MGANERFFLSCDINSVPTFRLEKIDGLPYKSSDSGVSIAE EKRPELYECALYIDAPFGLPWRTRKLTGPPYCMNELITLSKYRDLTAHSAQAIT VMVYSGKTEGLIGATVILFENKMGKSGKRLRMQGEKDGSEFSTGKPYRHE RGELEERLEKLMNTERGOIOSIDMLRLMKSLIDITKEOESYKHGSSHLFVYIDCSF EHRVYFQESGANFITAPIGSTNEEYVTDTEGRTNPSNKQLARSLDRGIIIDRL IKSPNTERKISQRYLVKPPRRTISGDRDRLKRPPLMSERBALFKLRQVWSDVO EAKOAIOLMKRMIDVCDMLSLFSESEYRAYAVSLRADDEELCTILOVO ALRERSDRSLQFLVQRLQNIETASLRLMYVAELHDYARFSTYELLENI IKLPVGNEDGYQWQSLVROQIETLAQCSITREVRNHYGRTQKRIKRLQGLGL SEULEEPEERPSLPYVNLKGVNAGESLFLKSLAPLPLPRTPEEGSGCLIRKKG DDLRODOLVQWVWMLDRLLKLENDLCLTPKPVATLADDEMLEFISRSQAOLISE HRSITSTYLOKEHPEDEHAPGITATCLDTEFKGAGSVTYTTLIGIDHLOVLLTD GRLEPHDFAITIGDRKPPRPBKIKKEVNAEGSGYTRRFGSCYCEANITIKS SNLIINFILMAGSTIPDIASDEKGIITLQEFRLMDDEACIFPDLLINESVAL FPMQVETIHRMAQYWR"
gene	15572..17812 /gene="F8A5.5" join(15572..16004,16236..17812) /gene="F8A5.5" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71966.1" /db_xref="GI:2462747" /translation="HWKKAATKKNSPSLAIKKRSRKNQDVSYEAPITSSYNDRI RPLDIYVRLNLMVNRGILPLIYVVGQSSKSVLSLGISLPRGQGCYVPL VMRLQSSSPSEPIEDVNDKVPYDDEHIAELIRATDYVAGSGSDAPLITLHV KACVPDITMDLDEGITRYVYVNGCPENITQISGMINEYREPOESITLVLSYVPTT CESIRKSRKYDKGQRTLAVYTFADNAPRGLQKVTADVSIVLGVCRNIGETTY EEAQMOBELFRTHPVLSLIDEDIVIGIYLAQKMLIQQSMARCLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDIIGASKSLAILVQGFSEYPPDOQNHCTARL ADMISQSDSKLQAKPREVAFLMEDEIKIIDECKCVGLPWFIRSFASLILQSHVQI QDKPEYFENKIMDYIEDVLSYFATKRSNDPFOOSIKRGRMLIKIKROQSNRNEI VEMEKLTDTYCNPEYMTSWTQSAOESFIDAVYKNEINTDYFSTVGGGNYISLNRK YHAHLIPADMKMRITSYKTYLRRITVYNALYQLSKSYLVNRFOKEIYAEAVYD RQGGCYEKLMEESPVLAASKREKLQNSIKILKESKQVAAIVDQC"
CDS	18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQANPHDNREVPYIEAPITSSYNDR IRLPLDYVRLNLMVNRGILPLIYVVGQSSKSVLSLGISLPRGQGCYVPL PIVRLQSSSPSEPIEDVNDKVPYDDEHIAELACATDYIAGGPISTYGRSSRY NSRPVNGCPENITQISGMINEYREPOESITLVLSYVATPHRRATQKQMOBELFR THPMLSLIDEDIVIGIYLAQKMLIQAATMISCLPITVIRKINHKETAVLELNKLPV MASTGEALMALMDIIGASKSLRILVQGFSEYPPDOQNHCTARLADLSQSDNLQ EKPEYEFELMEDEIKIIDECKCVGLPWFIRSFASLILQSHDAIVKVEYFIRKIMD YIEVLSVIYAKYSENPPQISAKIKRAGNRLITIKIHSYVERVLEIVEKELDTYCN

Db	3611	accatgccaatggaatgctgcttcgctgcacagagtgaaacaatcaatcaatcaaccc	3670
Db	71882	ACATGCCAATATGAAACCGTAGCTGATCCTCGAAGAGTTTCATTAACACATTCTCTGT	71823
Qy	3671	ctgctcctctcctcaattcactttgaacgtgaagtgcttttccatgcttctccttaeggtt	3730
Db	71822	TTTGTATATCTCTTCGATATATATACCGAAAGTTCTTATATAGATTT	71776
Qy	3731	cataattgtgtgttaacctaagaaacagagagcgcaaccgfcacaaactccgcttatgt	3790
Db	71776	-----ACATGTTTCAGACATATTAAGAGAAAGCCGTCTTAAGATTG	71733
Qy	3791	gccaaacggaagaagatcgcgctcagctcagctcgaagttgtgcttaccatcatgatca	3850
Db	71734	GAATCGGAGGATATAGGATTGCACTCGCGGACAGCTCGAGAGCTGTTATCTTTCACGAGCA	71675
Qy	3851	ctcgcatcggaagatcatctccaccgtgaagctaaagcagcaactcctcttaagagaga	3910
Db	71674	ATGGAAATCCAAATATTATTCACAGAGAGCTCAAACTCCAAATTTTACTTGAATGAGAG	71615
Qy	3911	attcgaagcggttgcttggaagattcgggttgcgaagtaatgacatataagaacactca	3970
Db	71614	CTTTGAAGCAATAGTTGGGAGATTTTGTCTACGAAAGCTTTTAACACAGAGAGTTCAACA	71555
Qy	3971	ctgtgaacaacagcagtcctgtgcacatccgttcaatcgtctccagaatatctctcaacgg	4030
Db	71554	TGTCATCATACCGCAGTCGAGGAACCATGTGGACATCCCTCCAGATACCTTTCCACTGG	71495
Qy	4031	aaactctcagagaacacgcaagtttctgcatcagaatcagtcagttcttctaga	4081
Db	71494	ACAGTCTCTCAGAGAAACCGATGTTTTCGGATTCCGAGTACTAATCTCTTGA	71444
RESULT	10		
LOCUS	AB013395/c		
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MON23, complete sequence.		
ACCESSION	AB013395	86064 bp	DNA
VERSION	AB013395.1	GI:3128142	
KEYWORDS	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MON23.		
SOURCE	ORGANISM	Arabidopsis thaliana	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euembryophytes; Spermatophyta; Magnoliophyta; eudicotyledons; cecidicots; Rosidae; eutosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE		1 (sites)	
AUTHORS		Nakamura, Y.	
JOURNAL		Structural Analysis of Arabidopsis thaliana Chromosome 5. VI	
REFERENCE		2 (bases 1 to 86064)	
AUTHORS		Nakamura, Y.	
JOURNAL		Direct Submision	
TITLE		Submitted (06-MAY-1998) to the DDBJ/EMBL/GenBank databases.	
URL		Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2, 1532-3, Yana, Kisarazu, Chiba 292, Japan	
FEATURES		(E-mail: ynakamura@kazusa.or.jp, Tel:+81-438-52-3935, Fax:+81-438-52-3934)	
source		Location/Qualifiers	
		1. 86064	
		/organism="Arabidopsis thaliana"	
		/strain="Columbia"	
		/db_xref="taxon:3702"	
		/chromosome="5"	
		/clone="MON23"	
		/clone_lib="Mitsui P1"	
		27710 a 15313 c 14986 g 28055 t	
BASE COUNT			
ORIGIN			
Query Match		6.7%	Score 273.2; DB 7; Length 86064;

	Best Local Similarity	61.1%	Pred. No. 3,1e-42;	
	Matches	484;	Conservative	0; Mismatches 268; Indels 40; Gaps 1
QY 3320	aaatctctatgcaaccggaagaatccagaattccatctctggtgacgtctcaagagtttc	3349		
Db 74153	AAAGTTTGATTTAGAGTGAAGTGGACAGAAAGGATTTGCTTTGGACATGTGAAGATTTG	74094		
QY 3350	ctttgcggagagccacaaagtcgagtgatgagtttagtaacaagaacatttggcgagag	3409		
Db 74093	CATGGAAGAGAGCTTCAAGTTGGCTACAGAGAGAGTTCAAGTGAAGAAAGATGTTCTCGACAG	74034		
QY 3410	gttggtttgggaaagtctcaagaagagcgcttgcgacgcgaaacctgtgtctgtcaaga	3469		
Db 74033	GAGCGCTTTGGGAAGGTTTACAAAGGATTTGCTTTCCGATGGCAGCAAAAGTCTGTAAAA	73974		
QY 3470	gacgcgaagagaagcggaacctccgagtgtagagctccagttccaaacgaagttagaga	3529		
Db 73973	GATTGACTGATTTTGAACGTCCAGGAGAGAGATGAAGCTTTCCAGAGAGAGATTAGATGA	73914		
QY 3530	taagatgagagatctcgaanaacctgttgaagattacgaagttctgtatgacacgagcg	3589		
Db 73913	TAAGTAGCTGTTCAATAGGAATCTCTTCGCCCTTATCGCTTTGTGTACACAAACATG	73854		
QY 3590	agagattgcgttgtatctctctacatgagccaagaagagtggttcgtgctcgagagta	3649		
Db 73853	AACGACTTTTGGTGTATCTTTTATCAGACAAATCAAGTGTGCATATGCTTAAGAGATA	73794		
QY 3650	aaaactaaacaataaacatcctgtgctctctcctaactcttgaacgtgaagtgtttt	3709		
Db 73793	AAAAAAGGACATTA-----GCTAT	73774		
QY 3710	tcatgtttcccttaatggttcatattgttgttaacctaatgacaagaagagccacc	3769		
Db 73773	AAACTTTTGATTTCAAGAGTTTAAAGTTTGGATCTATATCAATGACAGAGATTAAAC	73714		
QY 3770	gtcaaacctccgcgttattgagccaacgcggaagaagaatcgcgcctagctcaagtcag	3829		
Db 73713	CGGGATTCAGTCTGTGATTTGTTTCAAGAGGAAACAGATGCGTTAGTGCAGCAGAG	73654		
QY 3830	tttgcttaccctacatgatactgcgactcgaagatcattcaacgcgtgaacgtaaagcagc	3869		
Db 73653	ACTGGAATATCTTCATGATGACATTCGCAACCCGGAAGATCATACAGAGATGTGAAGGTGC	73594		
QY 3890	aaacatctcttgaacgaagaattcgaagcgttttcttgaagattccgggttggagaagct	3949		
Db 73593	AAAGTGTACTATATCAAAACCTTTGAAGAGATGTTGTTGATTTTGGTTTAACCAAGTT	73534		
QY 3950	aatggaactaataagaacatcagctgacaaacagcagtcgtygcacatcogtltacatgc	4009		
Db 73533	GGTAGATGTTGAAGAGCATATGATTAACCACTCAGGCTCAGAGAACAAATGGCTATATTCG	73474		
QY 4010	tccgaataatctctcaaccgcgaaatcttcagagaagaacagcgctttcggatcggaat	4069		
Db 73473	ACCGAAATGATATCCACAGGAAATCGTCAGAGAAACCGATGTTTCGGGTACGGGAT	73414		
QY 4070	catgcttctaga	4081		
Db 73413	TATGCTTCTGSA	73402		
RESULT 11				
LOCUS DCU93048	1755 bp	mrna	PLN	28-JUN-1997
DEFINITION	Daucus carota somatic embryogenesis receptor-like kinase mRNA,			
ACCESSION	U93048			
VERSION	U93048.1			
KEYWORDS	GI:2224910			
SOURCE				
ORGANISM	carrot.			
	Daucus carota			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
	eudicot; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.			


```

Db      821 AGGTGTATAGAGACGCTTGTCTGATGCTCATCTGTACGACTTAAAGCCTTAAAGAG 880
QY      3482 agcgaactccaggttgagagagccagcttccaaacagaagtagagataagatgagcag 3541
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      881 AACGAACACCAGAGTGTGAGCTGCACTTCAACAGAGAGTGAATGATTAGCAATGCGCTG 940
QY      3542 ttcactgaacacctgtttagattacagaggttctgtatgacacagcagagagatgtcttg 3601
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      941 TGCATGGAATCTTTCGCTCTACGCTGCTTCTGATGACACCAACAGCGGCTTCTTG 1000
QY      3602 tgcattccactacagcagcagtagaagtggtctgtctgtctcagaggttaaataacaa 3661
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1001 TAAATCCATACATGCTTAATGGAAGTGTTCCTGCTGTTTAAAG----- 1044
QY      3662 ttaaacatctgtgtctctctcctaactacttgcagtggaagtgttttctatgttctct 3721
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1044 ----- 1044
QY      3722 ttatggttcataatgttggttacctaataigacacagagagccaccgtcaaacctcc 3781
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1044 -----AGAGCGTCAGCATCAGAACCTTC 1067
QY      3782 gcttgaatggccacgcggaagaagaatcgcgtacgctcagctcgaaggttgccttacct 3841
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1068 CTTGATTTGGCCAACTAGAGAGAGATGCTACTAGAGATGCTAGGGGCTTCTTATT 1127
QY      3842 acatgatcaactgcgatcgaagaatcatcaccgtgaagcgttaaaagcagaacatctctt 3901
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1128 GCATGACCAATGTGATCCCAAGATTTTCATGCTGATGTAAGTGCAAAATTTTAT 1187
QY      3902 agacgaagaatcgaagcgtgtgttgagattcggattcggatggcgaagaatggaactaa 3961
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1188 GAGCAAGATTTGAGGCTGTCTTAGATTTTGGGTTAGCTAGGCTCATAGATTACAA 1247
QY      3962 agacactcagtgacaacagcagctccgtgacacatcggatcagatcgcagaatctc 4021
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1248 GGATACCCATGTACACGCTGTAAAGGGGTACTCTGGGCTCAAGACCTCCGACTACT 1307
QY      4022 ctcaacgcggaatactctcagagaacacgacttttcggtatcgaaatcgtctctga 4081
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1308 CTCGACTGGAAATCATCAGAGAACGATGCTTGTGTTATGGATTATATCTTAA 1367

```

RESULT 13
AB025639/c
LOCUS
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, pl clone: MML2, complete sequence.
ACCESSION AB025639
VERSION AB025639.1 GI:4589445
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (strain: Columbia) DNA, clone: MML2, clone: MML2.
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (sites)
AUTHORS Nakamura, Y.
TITLE Structural Analysis of Arabidopsis thaliana Chromosome 3. II
REFERENCE Unpublished (1999)
AUTHORS 2 (bases 1 to 84896)
TITLE Direct Submision
JOURNAL Submitted (02-APR-1999) to the DDBJ/EMBL/GenBank databases. Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2, 1532-3, Yana, Kisarazu, Chiba 292, Japan (E-mail: ynakamura@kazusa.or.jp, Tel: +81-438-52-3935, Fax: +81-438-52-3934)
FEATURES
 source
 1. 84896
 Location/Qualifiers
 /organism="Arabidopsis thaliana"

```

BASE COUNT      27260 a 15033 c 15617 g 26980 t
ORIGIN
Query Match      6.4%; Score 263.2; DB 7; Length 84896;
Best Local Similarity 61.6%; Pred. No. 2.4e-40;
Matches 483; Conservative 0; Mismatches 253; Indels 48; Gaps 2;
      /strain="Columbia"
      /db_xref="taxon:3702"
      /chromosome="3"
      /clone="MML2"
      /clone_1id="Mitsui P1"
      /clone_2id="Mitsui P1"
QY      3298 tatgcagccgaagaagatccagaagttcatctggagacgctcaagaggttcttcgcg 3357
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67679 TACACAGACCAAAACAGAGAGAAATGTCTAGGAACTACAGAGAGTAAATTTCMAA 67620
QY      3358 gagctcaaggttgagagtgatggtttagtaacagaacatttggcagaggtgggtt 3417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67619 GAACTTCAATCCGCACTGTACTTCAAGCAACAGAAATCTGTCGGAAGAGAGGCTTT 67560
QY      3418 gggaaagttcaagaggagccttgcagacggaactctgttgcgttcaagaagactga 3477
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67559 GGAATGTGTAAAGAGTGTCTTCATGATGAAATATCATCGCGTGAAGATTAAG 67500
QY      3478 gaagagcgaactccaggttgagagagctccaagttcaacagaagtagagatgataatg 3537
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67499 GATTTAAACAAATGTGTGAGAGATGTTCAAGTTCAGACAGCTTGAAATATAGCCTT 67440
QY      3538 gcagttcacgaacacctgttgatattacagaggttctgtatgacacgacccagagattg 3597
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67439 GCGGTCCACCGAGATCTCCGCTTATACGCTTCTGTACTACTCTCTGAACGGCTT 67380
QY      3598 ctgtgtatcccttaccatgagcgaatggaagtgtctgtctcagagaggttaaaactaa 3657
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67379 CTCGTTATCTCTTACATGTCATGCAATGCAATGCAATGCAATGCAATGCAAT 67320
QY      3658 acaattaaacatctgtgtctctcctaacttcttgcgtgaaggtgttttcatgtt 3717
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67319 AACA-----CCTTGATTTTGAATTTCTTACTAGATGATGACCGGT 67279
QY      3718 tccattatggttcataatgttgttgaactaatgacacagagagccacgcacaaac 3777
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67278 TACGATTCGTGTGAATGATGTTGCAGCTA-----AAC 67248
QY      3778 ctccgcttattggcccaagcggaaagagaatcgcgctagctcagctcgaaggttgcctt 3837
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67247 CGGATTTGATTTGGGCAACAAAGAACCAATAGCATTTAGAGACGAAGAGGTTGCTGT 67188
QY      3838 accatcatgatcatctgcgtccgaagatcatcaccgtgacgtgaacgaacacacacac 3897
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67187 ATTTGATGAGCAATGTGATCCAAAGATCATTCACCGATGATGATGATGATGATGAT 67128
QY      3898 tcttaagcgaagaatcgcgaagcgtgtgttgagatttgggttggcgaagtaatgtgact 3957
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67127 TTTCTTACGATTTCTTTGAAGCTGTGCGGATTTTGGGTTGGCTTGAAGCTTTTGATC 67068
QY      3958 ataaagacatcagtcgtgaacacagcagtcggtgcacacatcggttcacatcgctcag 4017
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67067 ATGAGGATCGGATGTGACACACCGCGGTGAGAGAAAGTGGTGTACATTCACACTGAGT 67008
QY      4018 atcttcaacgcgaataatctcagagaacacgcgaagtttggataggaatcatgtctc 4077
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      67007 ATCTTCAACAGCAATCTTGTGAGAACAGATGTGTGCTTGGGATTTCTTCTTC 66948
QY      4078 taga 4081
      | | |
Db      66947 TCGA 66944

```

RESULT 14
AC005170/c
LOCUS AC005170 110211 bp DNA PLN 17-DEC-1999

DEFINITION Arabidopsis thaliana chromosome II section 135 of 255 of the complete sequence.
ACCESSION AC005170 AE002093
VERSION AC005170.2 GI:6598448
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 110211)
AUTHORS Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D., Fujii,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E., Feldblum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.U., Ronning,C.M., Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Unayam,L., Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H., Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D., Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and Venter,J.C.
TITLE Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
JOURNAL Nature 402 (6763), 761-768 (1999)
MEDLINE 20083487
REFERENCE 2 (bases 1 to 110211)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT On Dec 17, 1999 this sequence version replaced gi:3738313. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (<http://www.tigr.org/tdb/at.html>).

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://arthur.epm.ornl.gov/pub/xgrail>), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://gnomic.stanford.edu/GENSCAN.html>), and NetPlantgene (<http://www.cds.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RX/repeatmasker.html>). Genes are numbered from the top to the bottom of the chromosome.

We thank the CSHL/Mashu/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Likin Zhou, Hanif Khailak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: atetigr.org.
Location/Qualifiers
1. 110211
/organism="Arabidopsis thaliana"
/cultivar="Columbia"

/db_xref="taxon:3702"
/chromosome="2"
complement(1..2793)
/note="Sequence from clone F217L4"
complement(356..401)
/rpt_family="(TAA)n"
complement(join(1549..2058,2731..3012))
/gene="At2g23810"
complement(1549..3012)
/gene="At2g23810"
/note="F217L4.1"
complement(join(1744..2058,2731..3003))
/gene="At2g23810"
/codon_start=1
/product="similar to senescence-associated protein"
/protein_id="AA18611.1"
/db_xref="GI:6598449"
/translation="MFLILVLCITPAFVVTNKGAGEALEGRGYREYKIGDYSTWL
QRRVNGKNMKIRSLVESKYSKLEAFVNPVNSFYKSHLALQSGCKPDECG
FHYVPTWTNKTGTHTNPDCOTDMAKKEKLCFQCCKAGLLDNVKSAMKVAIVN
IYFLVFLIVSYVGCACFRNNKRDSYRITGYRP"
2794..110211
/note="Sequence from clone T29E15"
complement(3567..3677)
/rpt_family="(TAA)n"
complement(3570..3641)
/rpt_family="(TAAA)n"
complement(4565..4612)
/rpt_family="(CAA)n"
6379..6484
/rpt_family="(CAAAA)n"
join(<6480..6749,7602..7726,7811..7848,7929..8039,
8152..8268,8487..8786)
/gene="At2g23820"
<6480..8786
/gene="At2g23820"
/note="T29E15.2"
join(6480..6749,7602..7726,7811..7848,7929..8039,
8152..8268,8487..8557)
/gene="At2g23820"
/note="unknown protein"
/codon_start=1
/protein_id="AAC63656.1"
/db_xref="GI:3738315"
/translation="MKNVAVFECRLSLVFLVPSFTSRSHIRFYSAAGASPNNRAH
CAASDPOSQDGSPPVAAVPPSSSSASAIIDFLSCTRLKTPRRGKIRDY
KDEPSTADHMYRMGALISDIPVNRDKCKMAIVDAIVGDIITPGGSKKE
KNRSEDALEHMKLIGSGERAKETAEIMRETEENSSPAKYVDFDVELLQALEY
EOKPQNTNGKMASETVSRKKQH"
complement(6654..6712)
/rpt_family="(GAA)n"
join(<9531..9582,9709..9779,9866..9947,10054..10169,
10268..>10396)
/gene="At2g23830"
<9531..>10396
/gene="At2g23830"
/note="T29E15.3"
join(9531..9582,9709..9779,9866..9947,10054..10169,
10268..10396)
/gene="At2g23830"
/note="unknown protein"
/codon_start=1
/protein_id="AAC63657.1"
/db_xref="GI:3738316"
/translation="MSNNELIEPMYLOPFPELKQMSCLYLTKNDNNVAFVKY
TNRNRCVBPNGYLIPKSTCKVLTVMQAKVPSDMOSFEKFIQSVYASPGYAKKE
VTRBMSKSGHVEETKLRVYVCSITTNITSSPTRNGFIQGC"
complement(11332..11371)
/rpt_family="(TAA)n"
15801..15917
/rpt_family="(TAAA)n"
15853..15918
/rpt_family="(CAAAA)n"

	repeat_region	16051..16136	/rpt_family="(TAAA)n"	
	mRNA	complement(join<20442..20537,20617..20691,20894..20948,21145..>21485)	/gene="Atg23840"	
	gene	complement(<20442..>21485)	/gene="Atg23840"	
	CDS	complement(join(20442..20537,20617..20691,20894..20948,21145..21485))	/note="T29E15.4"	
	gene	/pseudo	/note="unknown protein"	
	mRNA	join(<22999..23070,23189..23296,23402..>23554)	/codon_start=-1	
	gene	/gene="Atg23860"	/protein_id="AAC63658.1"	
	CDS	<22999..>23554	/db_xref="GI:3738317"	
		/gene="Atg23860"	/translacion="MAHFSARLKLNSNDGLSFVDSHDQFROSLYAGGSGPLVIPRMSRLKFCARIFSSSYSPFYRRKKRIDLDYMERKLLIDEDNDEMDFGDDDLIEDHLSCERLVLYDIRPVYCWKRKAICLEVMDKADVLVEYDQVFPSSTGSFYLPATLRAPHLLGVRRRYRNLSLRKNILLRDDYTQC"	
	gene	complement(22086..22781)	LLOVRRRYRNLSLRKNILLRDDYTQC	
		/gene="Atg22850"	complement(22086..22781)	
		/note="T29E15.5"		
		/pseudo		
	mRNA	join(<22999..23070,23189..23296,23402..>23554)	/gene="Atg23860"	
	gene	<22999..>23554	/gene="Atg23860"	
	CDS	join(22999..23070,23189..23296,23402..23554)	/note="T29E15.6"	
		/gene="Atg23860"	/codon_start=1	
		/note="unknown protein"	/protein_id="AAC63659.1"	
		/db_xref="GI:3738318"	/translacion="MTTNPKNYCVRPNGILIPKSTCKLRKPSPDMOSNEKEFIORVKASPAATAKEVTLETFNKESGHIVEIKRTIYYCSTTTIIISPRTRNGLFENFVRHRFTSL"	
	repeat_region	27493..27571	/rpt_family="(TAAA)n"	
	repeat_region	complement(29076..29146)	/rpt_family="POLYA"	
	gene	29302..29885	/gene="Atg23870"	
		/note="T29E15.7"	/pseudo	
	repeat_region	31472..31519	/rpt_family="(TA)n"	
	mRNA	complement(join<31877..32176,32267..32407,32498..32893,33162..35713,35886..>36147)	/gene="Atg23880"	
	gene	complement(<31877..>36147)	/gene="Atg23880"	
		/note="T29E15.8"		
	Query Match	6.28; Score 252.4; DB 50; Length 110211;		
	Best Local Similarity	58.98; Pred. No. 2.5e-38;		
	Matches 516; Conservative	0; Mismatch 316; Indels 44; Gaps 3;		
Oy	3209	ttaggaacactttaaatacaatcatcattgcttttaagtacaacctc	3268	
Db	54432	TTCGCATTAGTGCGTTAAGCAATCAGAGAAGCTTAGTAAATTGAATTGTGTTT	54373	
Oy	3269	tttgagtaaaacttcgatcgcaaatctctatgcacggagaagctcag--aaagtc	3325	
Db	54372	GTTGCTAAGATATTTTGTGGTCTTAAATGACAAAGCAAGCAAGGAAGGATTACTTG	54313	
Oy	3326	atctggagacagctcaagggtctctcttcgsgagatcacaaagtgcagtgatggttta	3385	
Db	54312	GTTGGGAATCATGAAGAAGCTTACATTCACAGGAACTTCATGAGTACGATGCTTTTA	54253	
Oy	3386	gtaacagaacacttttggcagaggtggttttgggaagctcacaaggagacgttcgacg	3445	
Db	54252	GTTCCAAAGATATCTCTGGCTGCTGGTGGTTGTTAATGCTACAGAGAAATTCGGGG	54193	

QY	3446	accgaatctctgttctgtcagaagacctaagaagagcgacatccagtgagagctcc	3505
Db	54192	atgggacagatgggttcgaagtcgaacgatttgaagaatgtgaatgaagacctccgggaactcac	54133
QY	3506	agtttcaacagagaagtagagatgatataagtaatgacgttcacatgaacctgttgagttac	3565
Db	54132	agtttctactatgacgttgatgagatgatacgtttgactgttcattggaatttgcctgggttaa	54073
QY	3566	gaggttctctatgaacacgcagcaggaagcttctgtatctactactacgtgcgaatgaa	3625
Db	54072	tcgggttatttgcgagttctacgcgaagaacctcttctttaccctttacatgtccatgcca	54013
QY	3626	gtgttgcttcgtgtctcagaagyttaaacaatacaataaatacatctgtgtctctctca	3685
Db	54012	gcgcgcgcctctagacgtccaaagagtcgcaaaaaaaggtttcgtatgatactgctttagtaa	53953
QY	3686	atactcttgacggaagtggttttcaatgtttcccttaatgatgttcataatgttggtta	3745
Db	53952	aaaa---aaataatgaatcacttgaaactatgttttggtaactgttca-----	53508
QY	3746	cactaatgacacagagaagccaacctgcacaaacctccgctgtatgtgcacaacgcggaagag	3805
Db	53908	-----gctaaagccagcgttgacactggaacacaaagagaa	53874
QY	3806	aatcgcgctatgctcagctcagctcgaagttgtcttactactactatgatactcgtcgatccgaagat	3865
Db	53873	gattgcacattggagactgcacagagaggttgattttatctacacgcagacatgcgatacccaagat	53814
QY	3866	cattcacgcgagacgtlaaaagcagacaacacatccctctgaacgagaatctcgagcggttgt	3925
Db	53813	tattcacgcagagatgcgaagcagacaaatcttcctcagatgagatgtttggaagcagttgt	53754
QY	3926	tggagatcttgcgggttggcgaagtaatgtaatgtaacataaagaacatcagctgacaacagcgt	3985
Db	53753	tggggatttttggactgtgacaaagcttactccaaacgaagagatttcacatgttcacaaacccgct	53694
QY	3986	ccgctgagacacatcggtacacatcgtccagaatactctcacaacggagaatcttcagagaa	4045
Db	53693	tagaggaactgttggtcactacgttcacactcgatgatactcaccacggtcagtcattcgagaa	53634
QY	4046	aaccgacgcttgcgatacgaatactatgctcttaga	4081
Db	53633	aaccgactgttgggttcggtactgttttgcctaga	53598
RESULT 15			
AB005234			
LOCUS	AB005234	89779 bp	DNA
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MBK5, complete sequence.		
ACCESSION	AB005234		
VERSION	AB005234.1	GI:2264306	
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_1lb:Mitsui PL clone:MBK5.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1 (sites)		
AUTHORS	Sato,S., Kotani,H., Nakamura,Y., Kaneko,T., Asamizu,E., Fukami,M.,...		
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned pl clones		
JOURNAL	DNA Res. 4 (3), 215-230 (1997)		
MEDLINE	97471969		
REFERENCE	2 (bases 1 to 89779)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUL-1997) to the DDBJ/EMBL/Genbank databases.		

Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of
Gene Structure 2: 1532-3, Yana, Kisarazu, Chiba 292, Japan
(E-mail: ynakamura@kazusa.or.jp, Tel: +81-438-52-3935,
Fax: +81-438-52-3934)

FEATURES

Location/Qualifiers
1. 89779
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MBK5"
/clone_lib="Mitsui P1"

BASE COUNT 28415 a 16340 c 16159 g 28865 t
ORIGIN

Query Match 6.0%; Score 244.6; DB 7; Length 89779;

Best Local Similarity 59.8%; Pred. No. 7,6e-37;

Matches 468; Conservative 0; Mismatches 279; Indels 36; Gaps 2;

```
QY 3299 atgcagccggaagaatccagaagttcattcgtgacagctggaagaggtttcttgcgag 3358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54502 ATGCAAGGGAAGATGACGAGAAATTCTTGTGCACTAAGCAATCTCTTACGTG 54561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3359 agctacaagtggcagttgattggttaacaagaacatttggcagaggtgggttg 3418
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54562 AATCCAGCTCCGACACGATATTTCACGAGAGCAATTTGATAGACAGAGGATTG 54621
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3419 ggaagttacaaggaagcgttggcagcgaactcttctgtcgaagactgaagg 3478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54622 GTAAAGTATACAGAGTTTGCTCCAGCAAAACAAAGTTGACGTGAACGCTTGCG 54681
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3479 aagagcgaaactccaggttgcagctccagttccaacagaagtagatgaatag 3538
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54682 ATTACTTCAGTCTTGAGAGAGAGCTGCTTCCAAAGAGAGATTCAGCTCAAGCGTTG 54741
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3539 cagttacgcgaacctgttgagattcgaagttctgtatgacaccgacgaagattgc 3598
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54742 CGGTTCATTAATCTTACGCTTATGGCTTGACAACTTCCTCGAAGAAATCC 54801
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3599 ttgtgatccttaccatggccaatgggaagtgttgcgttcgttcagaggtaaaaactaa 3658
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54802 TTGTTATCCATACAGAAATCTTAGTGTGACATATCGACTAAGAGGTATATA---- 54857
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3659 caatlaaacctctgtcctcctcaattacttgacgtgaagtttttccatgttt 3718
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54857 -----TACTTAGCTCTCTATACATTACCAATAATAGATTCTATATGTGTG 54907
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3719 ccttaatggttcataattgttggtacataatgacacagaagggccacgctcacacc 3778
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54908 TACATATGTG-----ACATTGTAGATTGAAAGCGGAGAGGA 54945
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3779 tcgccttgattggccaagcggaagaatcgcgtaagctcagctcgaggtttgtccta 3838
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54946 AGGATTAGACTGGCCAAAGAGCGGTAGCTTTTGGTTCAGCTCAGCGTTTAGAGTA 55005
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3839 cctacatgctacgtcgaatccgaagatcattcacgctgaacgttaaaagccaaacacct 3898
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55006 TCTACACGACATTTGTAACCCGAGATCATACCGGATCTCAAGCTGCAAACTACT 55065
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3899 cttaagcagaagaatcgaagcgggttggtggaattcgggttgccaagctaatagacta 3958
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55066 TTTAGCAACAATTTTGAGCGAGTTCTTGAGATTTCGGTTTAGCTAAGCTTGTTGACAC 55125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3959 taaagcactcactcagatgacaacagcagtcgttgccacatcgctcacatccagaata 4018
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55126 ATCTCTGACTCATGTCAACAATCGAGTCCAGGACCAATGGTCACTTGGCCAGAGTA 55185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4019 tctctcaacccggaanaatcttcagagaanaaccgacgttttcggaatacgaatcgttct 4078
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55186 TCTCTGACACGAAATCATCTGAATAAACCGATGTTTGTGTTACGATTAACGCTTCT 55245
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4079 aga 4081
```

Db 55246 TGA 55248

Search completed: June 24, 2000, 00:12:15
Job time: 45690 sec

QY 1921 ataagagtgtaagcttctcttactaatcccactttttaacttgcctcagcgtggt 1980
DB 1921 AATAAGAGTGAAGCTTCTCTACTAATCCACATTTTAACTTGAACGAGGTG 1980
QY 1981 tacgcgaatcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2040
DB 1981 TACCGACATTTTGTCTTGTGTGCAAAATACAGATTTGGGAATGACAGATTCGG 2040
QY 2041 ccatttagttccagaagcttggtgtgtcgaagaattgcagtaattgttaagttccatt 2100
DB 2041 CCATTTAGTTCAGAGCTTGTGTGTCTCAGAGATTTGCAATTTGTGTAAGTTCCACTTAT 2100
QY 2101 gcatcatgctttaacaaaacaaatccaaagatttgacagaagaagcctgagttacctt 2160
DB 2101 GCATCATGCTTTAAACAAAACAAATCCAGATTTGACAGAGAAAGCACTGGATTCCTTT 2160
QY 2161 tctaatgtgaatcctttttaacaaagttcttattcttcttaccagggagcttaccagtaaca 2220
DB 2161 TGTAAATGAAATCTTTTAAACAAAGTTCTTATTTCTTACAGGGAGCTTTACAGTAAACAA 2220
QY 2221 cataactggcccgatcttcctagtaactcttggaacttgacaacattagtgagtttgact 2280
DB 2221 CATTAATGCCCCGATTCCTCAGTAATCTTGGAATCTGACAAACTTATGAGTTTGATCT 2280
QY 2281 ttaacttaaacagctctcccggtccctactccggaaatcattggygaagccttcaagctgag 2340
DB 2281 TTACTTAAACAGCTTCTCCGGTCTATTCGGAATCATTTGGGAAAGCTTTCAAGCTGAG 2340
QY 2341 attctgtgagtatatacatatgctttaccggctcagttaccagctcttcttcttcttctt 2400
DB 2341 ATTTCTGTAGATACATATATGCTTTACCGGCTCAAGTTACAGTCTTGTATTAATCTTAAGT 2400
QY 2401 ttgtgtccaatcttggaactctgtcgtgaagaatttaccatgacaagaatagccggcttaaca 2460
DB 2401 TTTGTTCCAAATTTTGCATCTTGTCTGANAATTTTACATGCAAGATAGCCGGCTTAAACA 2460
QY 2461 acaaacagctcacggtgtcaatcttccatgltcacgtgacaataatatacctccctcaagt 2520
DB 2461 ACAACAGTCTCACTGGGTCAATTCCTATGTCACTGACCAATATTAATACCTTCAAGT 2520
QY 2521 tctgtgctctctcaataatcttcaattatgcttactcaattctccctcagttgattgtt 2580
DB 2521 TGTGAATCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2580
QY 2581 gagttaaagcacttaacacttgatgatagacacagagaatctataaataaagaactctc 2640
DB 2581 GAGTTAATGCACTTAACCTTGAATGATGACACACAGAGATCTAATAATTAACAGACTCTC 2640
QY 2641 tggttcagttcctgacaatgagcttccctcactcttcaacaccatcagggttctatgatt 2700
DB 2641 TGGTTCAATCCGACAAATGGCTCTTCACTTCACTTCAACACCAACAGGTTCTATGATTT 2700
QY 2701 atcccttcaagtattatctcagttgtgtgtcagtgctgaactatctcgaacttctcatt 2760
DB 2701 ATCCCTTCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
QY 2761 tccctgtgcaagtttgctataactagaactatgagacgtgagacgtgttaacagttaccatg 2820
DB 2761 TCCCTTGAAGTTTGTCTAATAACTTATGACCTATGAGCCCTGTTTACAAAGTAAACCATGT 2820
QY 2821 cctgagctcccccgtttctccctccacacacttbatccaactcccccagtttccacc 2880
DB 2821 CCGGAGTCTCCCGGTTTCTCCCTCCACACCTTTATTAATCAACGTCGCCCAAGTTTCCACC 2880
QY 2881 ccgagtaagacctctctcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2940
DB 2881 CCGAGTAAGCCTCTCTTTTAACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940
QY 2941 ctctgcaatccttcttctataactcatcttgcaataagggacataccaatgac 3000
DB 2941 CTCTGCAATCCCTTTTCTCATATACTCATCTTGCCAAATAAGGCAATAACCAATGATC 3000

QY 3001 taattgattccagtggtgtatgtatbaacttggagcaatagctgtgtgagttgtctcag 3060
DB 3001 TAATTGATTTCAGTGGGATGATGATTAACGTGAGCAATAGCTGGGATGCTGCAGG 3060
QY 3061 tgcgtcttgcctcttgcctcctccaaatagccttgccttgccttgccttgccttgcct 3120
DB 3061 TGCCTTGTGCTCTTGTGCTCTCTCAATAGCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
QY 3121 actagatattctctcagatgctgctgtgtatattatattcagatagttctgttctta 3180
DB 3121 ACTAGATATTCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
QY 3181 gccagcaatttcttcttgcagaagaatlttggaacaactgttaatgaanaatcaatacata 3240
DB 3181 GCCAGCAATTTTCTTGTGCAAAAGTATGGAACAACCTGTTAATGAATAATCAATACATA 3240
QY 3241 agtcattgtttttaaagttacaactcttlttgatgaatctcogattggaanaatctctat 3300
DB 3241 AGTCATTGTTTAAAGTTACAAACTCTTTGATGAATAATCTCGATTGCAAAATCTAT 3300
QY 3301 gccagcgaagaagatccagaagttcatcttggaacagctcaagagtttcttcttgcgag 3360
DB 3301 GCAGCCGAAGAAGATCCAGAAGTTCAATCTGGGACACTCAAGAGTTTCTTGGCGGAG 3360
QY 3361 ctacaagtgcgagtgatggttttagtaacaagaacatttggcgaggtgggttggg 3420
DB 3361 CTACAAATGCGCAGTGTATGAGTTAGTAAACAGAAACATTTTGGCGAGAGTGGGTTGGG 3420
QY 3421 aaagttcaagaaggagcgttggcagacgaactctgttgcctgttcaagagacttgaagaa 3480
DB 3421 AAAGTTCAAGGAGAGCCTTGGCAGACGGAACTTGTGTCTGTCTCAAGAGACTGAAGGAA 3480
QY 3481 gaggaaactccagtgtagagagccagtttcaacagaagtagatgaatgaatgaga 3540
DB 3481 GAGGAACTCCAGGTGAGAGCTCCAGTTCAACAGAACTAGATGATTAATGATGATGATGAT 3540
QY 3541 gtccatcgaaacctgtgagattacagagtttctgtatgacacgcagacagagatgctt 3600
DB 3541 GTTCATGGAACCTGTTGAGATATGAGAGTTCGTATGACACGACCGAGAGATGCTT 3600
QY 3601 gtgatcccttaactgagcaatggaagtgttgccttgcctcagaggttaaaactaaca 3660
DB 3601 GTGATCTTACATGAGGCAATGGAAGTGTGCTGTCTGAGAGTAAACCTAAACA 3660
QY 3661 attaaacatcttgctctcctccaaattacttgaagttgaagtttcttcaatgcttcc 3720
DB 3661 ATTAAACATCTTGTGCTCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3720
QY 3721 ttatgggttcaataatgttgtgttacaactaagacacagagagggccacggtcacaactc 3780
DB 3721 TTTATGGGTTCAATTAATGTTGTTTACACTATGACACAGAGAGGCCACGTCACAACTTC 3780
QY 3781 cgccttgattggccaacgggaagaagatcgcgctagagctcagtcagagtttgccttacc 3840
DB 3781 CGCTTGATTTGGCCAAACCGGAAGAGAAATCGGCTAGGCTCAAGCTCAAGGTTTGTCTTACC 3840
QY 3841 tacatgacactgagatccgaagaatcatcacccgttgaagttgaagcagaacaactcct 3900
DB 3841 TACATGATCACTGGAATCCGAAGATCATTAACCTGATGACGTAAGAAACGACAAATCTCT 3900
QY 3901 tagacgaagaatttcgaagcgggttgttgagatttcgggttgcgaagactaatgagactata 3960
DB 3901 TAGACGAAGAATTTGAAGCGGTTGTTGAGATTTGGGTTGGCAAAAGCTAATGAGACTATA 3960
QY 3961 aagacactcccggtgacaacagcagtcogtggacacatcogtgaacatcogtccagaatatt 4020
DB 3961 AAGACACTCACGTGACACAGCAAGCTCCGTGGCAACATCGGTCACATCGCTCCAGAAATATC 4020
QY 4021 tctcaacgggaanaattctcagagaanaacgagcttctcagatcgggaatcatgctctag 4080
DB 4021 TCTCAACGGGAATCTTCAAGAGAAAACGACGTTTTCGATACGGAATCATGCTCTTAG 4080
QY 4081 a 4081

Db 4081 A 4081

RESULT 2
V06591 standard; cDNA to mRNA; 2089 BP.

AC V06591.03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK gene.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FT Key location/Qualifiers
FT CDS 195..2072
FT /tag= a
FT /product= SERK protein

PN W09743427-A1.

PD 20-NOV-1997.

PF 13-MAY-1997: E02443.

PR 14-MAY-1996; GB-010044.

PA (NOVS) NOVARTIS AG.

PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;

DR WPI; 98-086529/08.

DR P-PSDB: M47023.

PT Production of apomictic seeds - useful in plant breeding

PS Claim 27; Pages 91-95; 123pp; English.

CC The sequence is that encoding SERK, a putative receptor kinase.

CC It may be used as part of a method of producing apomictic seeds

CC comprising: (a) transforming plant material with a nucleotide

CC sequence encoding a protein which in active form in a cell or

CC cell membrane renders the cell embryogenic; (b) regenerating

CC the transformed material into plants or carpel-containing

CC plant parts; and (c) expressing the sequence in the vicinity

CC of the embryo sac. The apomictic seeds and embryos thus produced

CC can be developed into plant progeny. This is useful in plant

CC breeding programs. Controllable and reproducible apomixis provides

CC many advantages in plant improvement and cultivar development in

CC the case that sexual plants are available as crosses with the

CC apomictic plant. Apomixis provides for true-breeding, seed

CC propagated hybrids and could shorten and simplify the breeding

CC process so that selfing and progeny testing to produce and/or

CC stabilise a desirable gene combination could be eliminated.

CC Apomixis allows plant breeders to develop cultivars with

CC specific stable traits for such characteristics as height,

CC seed and forage quality and maturity.

CC Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;

Query Match

Best Local Similarity 13.2%; Score 539.2; DB 1; Length 2089;

Matches 664; Conservative 0; Mismatches 3; Indels 113; Gaps 1;

QY 3302 cagccgcaagaagaatcagaagtcctgagcagaactaagaagttcttgcggagc 3361

DB 1018 CTGCCGAAGAAGTCCAGAACTTCATCTGGACACTCAAGAGGTTTCTTGGCGAGC 1077

QY 3362 tacaaatgacgaatgagtttaagtaacaagaacatttggcagagtggttgga 3421

DB 1078 TACAATGGCGAGTGTAGTAACAACAACATTGGCGAGAGTGGTTGGGA 1137

QY 3422 aagctacaaggaagccttgcagacgaactctgttgcgtcgaagaactgaagaag 3481

DB 1138 AAGTCTACAAAGGACCTTGGCAGACGAACTCTGTCTCAAGAGACTGAAGGAAG 1197

QY 3482 agcgaaactcaggtgagagctccagtttcaagaagaatgaagaatgaatgagag 3541

DB 1198 AGCGAACTCCAGGTGAGAGCTCCAGTTTCAACAAGAAATGATGATGAGCAG 1257

QY 3542 ttcatcgaacacctgttgaagattacagaggttctgtatgacacgcagcagagattgcttg 3601

DB 1258 TTCATCGAAACCTGTGAGATTACGAGTTCGTGATGACACGACGAGAGATTGCTTG 1317

QY 3602 tgatccctacatgagcgaatggaagtgtgtctgtctgtctagagtgtaaaactaaaca 3661

DB 1318 TGTATCTTACATGGCCAAATGGAAGTGTGCTGTGCTAG----- 1361

QY 3662 ttaaacatctgtgtcctctcctcaacttgaacttgagcgaagtgttttcaatgttcc 3721

DB 1361 ----- 1361

QY 3722 ttatggttcataattgttgttacaactatgacacagagagccacgttcacaactcc 3781

DB 1361 ----- -AGAGGCCACCGCCACCAACCTCC 1384

QY 3782 gcttgatggccaacgagagaagaatcgctagagctcagctcgaagttgtcttacct 3841

DB 1385 GCTTGATGGCCAAACCGGAAGAAATGCCCTAGGCTCAGCTGAGATTGTCTTACT 1444

QY 3842 acatgacactgacatccgaagatcatccagtgacgttaaaagcaacaactcctc 3901

DB 1445 ACATGATCAGCTGCATCCGAAGATCATTCACCGTAGCTAATAAGCAACAACCTCTT 1504

QY 3902 agacgaagaatccgaagcgtgtgttgagagatttcggtgttgcaagctaatgagctata 3961

DB 1505 AGACGAAGAAATTCGAGCGGTTGGAGATTTCGGGTTGGCAAAAGCTTATGACATATA 1364

QY 3962 agacaactcagtgacaagaacagcagtcggtgacacatcgctccaatatct 4021

DB 1565 AGACACTCAGCTGCACAAACAGCAGTCCGTGGCACCATCGGTCACTCCCTCCAGAAATCT 1624

QY 4022 ctcaacccgaaatcttaagagaanaacgacgcttctggatacgaatcatgctctaga 4081

DB 1625 CTCACCGGAATCTTGAGAGAAACCGACGATTTTCGATCGGAATCATGCTTCTAGA 1684

RESULT 3

V06570 standard; DNA; 6695 BP.

AC V06570.03-AUG-1998 (first entry)

DE Daucus carota SERK gene.

KW receptor kinase; apomixis; apomictic; seeds; production; embryos;

OS Daucus carota.

FT Key location/Qualifiers

FT CDS 3696..6620

FT /tag= a

FT /note= "contains introns"

FT /number= 3802

FT /tag= b

FT /number= 1

FT /tag= c

FT /number= 3851..3979

FT /tag= c

FT /number= 2

FT /tag= d

FT /number= 4124..4211

FT /tag= d

FT /number= 3

FT /tag= e

FT /number= 4

FT /tag= e

FT /number= 4430..4528

FT /tag= f

FT /number= 5

FT /tag= g

FT /number= 4642..4757

FT /tag= g

FT /number= 6

FT /tag= h

FT /number= 4890..4967

FT /tag= h

FT /number= 5295..5803

FT /tag= i

FT /number= 6197..6339

FT /tag= i

FT /number= 9

PN MO9743427-A1.
PD 20-NOV-1997.
PR 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR MPI-98-086529/08.
DR P-ESDB; W47013.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 21; Pages 40-46; 123pp; English.
CC The sequence is that encoding SERK, a putative receptor kinase.
CC It may be used as part of a method of producing apomictic seeds
CC comprising: (a) transforming plant material with a nucleotide
CC sequence encoding a protein which in active form in a cell or
CC cell membrane renders the cell embryogenic; (b) regenerating
CC plant parts; and (c) expressing the sequence in the vicinity
CC of the embryo sac. The apomictic seeds and embryos thus produced
CC can be developed into plant progeny. This is useful in plant
CC breeding programs. Controllable and reproducible apomixis provides
CC many advantages in plant improvement and cultivar development in
CC the case that sexual plants are available as crosses with the
CC apomictic plant. Apomixis provides for true-breeding, seed
CC propagated hybrids and could shorten and simplify the breeding
CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match 9.5%; Score 387.4; DB 1; Length 6695;
Best Local Similarity 57.8%; Pred. No. 7.3e-78;
Matches 903; Conservative 0; Mismatches 506; Indels 152; Gaps 6;

QY 2197 ttacagagagcttaccagtaacaacatactgcccagctctagtaactcttggaactc 2256
DB 3974 TGACAGGAGCTTTACAGCAATATACATAGTGACCAATTCCTAGTATGGGAATCT 4033
QY 2257 gcaaaccttagtgagcttgatcttacttaaacagctctccgctctatccggatc 2316
DB 4034 GACAAATTTGGAGCTTGACCTATACATGATAGCTTCTGGACCTATACCGGACAC 4093
QY 2317 attgggaagcttcaaacctgagatctctgagatatacatatcttaccgctcagt 2376
DB 4094 ATTAGGAAGCTTACAAAGCTTAAGATTTCTGTATGACTACAAAT-----C 4138
QY 2377 tacagctcttgcttaactccttaggtttgtcccaattttgacctcttgcgaattta 2436
DB 4139 TTCACTAGTGTTTTAACTTATATGCAATTTGATATCTCTTTCAGTATGATTATATCA 4198
QY 2437 catgcagagatagcgagcttaacaacacagctctcactgaggtcgaattctctgtaactg 2496
DB 4199 AATTACTGATAGGCTCTCAACAACTGCTCTCTGTGTCATTTCCATATGTCACATG 4258
QY 2497 ccaattactaccccttcaagtggttgagctccctcattaaattcttaagtctact 2556
DB 4259 CTAAATTTCAACTCTTCAAGTCCGTGATTCGACCTTCCAGATAGT----- 4312
QY 2557 tcaatccctcagtgatttggttgagtaagcaacttaacttgatgagcaacagac 2616
DB 4312 -----TTGTTGTTGGATGTTTCAATTTTAAATCAATTAATGTTTCATCAG 4357
QY 2617 agactatacaataacagactctctgltcagltcctgcaaatggtcctctcactctt 2676
DB 4358 GGATTTCAAAACATCGGCTATACAGCACAGTACCGGATATGCTATTTCTTTTGT 4417
QY 2677 cacaccatcaggtctatagattatcctcctcagttatttcaagtggttgagtgagtc 2736
DB 4418 TACACCTATCAGGTTTAAAGTATGTAATCTTTAATATTAAGTCTTACTTCACTGC 4477
QY 2737 tgacttatctlgaaacttcatctc-----cttgcgagtttgctaa 2780

DB 4478 GAAACCTATGATTAATTTTCTCTCCATATATATATACCTTCGACGTTTGGCAA 4537
QY 2781 taacttagacctatgtagacctgttacaagtcaccacatgctcctgagatcccccgtttc 2840
DB 4538 TAATTTGAATTTAGTGTGACCTGTACTGGGAGGCCCTGGCCCTGAGATCCCATTTTC 4597
QY 2841 tctccaccctcttattcaacctccccagtttccaccgccgagtaagcctcctctt 2900
DB 4598 TCCACACACTCCGTTCAATCCACATCAACAGTACAGCTCCGAGTGATTAATTTTAT 4657
QY 2901 tagttacattataggaacagagatgaattcttctctctgcuaacccctttcc 2960
DB 4658 ATTAATTTCCCGTAATTAATTTTATGACGTAAATTTGGTGTATTTCCAGGTTGCCA 4717
QY 2961 atataactcattctgccaataaggcaataccaatgatcatttgatttcagtgagtt 3020
DB 4718 ATAAAGTATTTTCCCTTCTTT-----CTCTCTTATTAATTAAGAGCAAA 4763
QY 3021 atggtatactggagcaatagctgtagtgagttgctgagtgctgcttcttctgctg 3080
DB 4764 ATGTTCCACTGGACCTATTTCTGGGGAGTACGCTGGTGCTGCTTACCTGTTCGCTG 4823
QY 3081 cctccgcaataagccttctgctgtagcgagcaagaagccactagatattcttcagat 3140
DB 4824 CACCTGCAATGGCATTTGATGATGATGGGAGAGAAACCGGAGAACATTTCTTTGATG 4883
QY 3141 tgcctgtagatttatcttcgcagtaagttctcgttcttagccagaatttcttgca 3200
DB 4884 TCCAGGTTAGTCCGTGAATTAAGTA----- 4910
QY 3201 gaaaagtattggaacaactgtaataatcaatacataagtcattgttttaagta 3260
DB 4910 -----TCTATTAAGGCT 4923
QY 3261 caactctttagtaaaaaatcgtatgcaaaatcctatagcagcgaagaatccaga 3320
DB 4924 TACTGTCTGGAGCTTTTCTTCACTGATTAATTAATTAATTAATTAATTAATTAAT 4983
QY 3321 agttacatcggaagcagctcgaagagtttcttctgagagagtaagtgagtgatg 3380
DB 4984 AGTGCACCTTGCTCACTGAAAGAGGTTTCTCTCGAATTTGCAAGTCCGAAAGGATAC 5043
QY 3381 gtttagtaacaagaacatttgggcaagtgaggttgggaaagtctacaagagagcct 3440
DB 5044 TTTTGTATC-----CAGCTTGAAGAGTGTGATTTGTATAGGTGTATTAAGGACGCT 5097
QY 3441 ggcagacggaactctgtgtgtctcaagagactgaaggaagagcgaactccaggttgaga 3500
DB 5098 TGCTGATGCTCCTCACTGTAAGCAATTAAAGGCTTAAAGAAAGAACACACAGGTGGCA 5157
QY 3501 gctcagattcaacaagaagtagagatgataagtagtgcagttcatcgaaacctgttgag 3560
DB 5158 GCGTCACTTTCAAAAGAGAGTGAATATATAGATGCTGTGATGAATATCTTCTGCG 5217
QY 3561 attcagaggtctctgtagaaccgagccgagagatgcttgtgtagtcccttaacatgacaa 3620
DB 5218 TCTACCTGGTTTCTCATGACACCTTACCGAGCGGCTTCTTATATATCAATACATGCTAA 5277
QY 3621 tgggaagtgttgctgtgtgtcagagagtaaaaaactaaacaattaaactcttgagctc 3680
DB 5278 TGGAACTGTGGCTCATATGTTAAGAGTATCTCACTTCAATTTCAATTTCCAGAA 5337
QY 3681 tctcaattacttgaacgtagtgatttttcatglttctccttattggttataatgtt 3740
DB 5338 GTTGTGTATTAATAAATGAATATATACCTCCATACATATGTATTAAGGTATATATTC 5397
QY 3741 g 3741
DB 5398 G 5398

RESULT 4

V06571
ID V06571 standard; cDNA; 1814 BP.
AC V06571;
DR 03-AUG-1998 (first entry)
DE Daucus carota SERK gene.
KW Receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW Plant breeding; ss.
OS Daucus carota.
FH Key Location/Qualifiers
FT CDS 94..1755
FT /tag= a
FT /product= SERK protein
PN MO9743427-A1.
PD 20-NOV-1997.
PE 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARRIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR MPI: 98-086529/08.
DR P-FSDB; W47013.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 21; Pages 47-51; 123p; English.
CC The sequence is that encoding SERK, a putative receptor kinase.
CC It may be used as part of a method of producing apomictic seeds
CC comprising: (a) transforming plant material with a nucleotide
CC sequence encoding a protein which in active form in a cell or
CC cell membrane renders the cell embryogenic; (b) regenerating
CC the transformed material into plants or carpel-containing
CC plant parts; and (c) expressing the sequence in the vicinity
CC of the embryo sac. The apomictic seeds and embryos thus produced
CC can be developed into plant progeny. This is useful in plant
CC breeding programs. Controllable and reproducible apomixis provides
CC many advantages in plant improvement and cultivar development in
CC the case that sexual plants are available as crosses with the
CC apomictic plant. Apomixis provides for true-breeding, seed
CC propagated hybrids and could shorten and simplify the breeding
CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

Query Match 6.6%; Score 270.8; DB 1; Length 1814;
Best Local Similarity 64.6%; Pred. No. 7.6e-52;
Matches 304; Conservative 0; Mismatches 157; Indels 119; Gaps 2;

QY 3302 cagccgagaagaagatccagaagatcattcgtggacagctcaagaaggtttcttgcgggagc 3361
DB 707 CAGCTGAAGAGGAGCCAGAGAGTGACCTGTGTAAGAGAGGTTTCTCTGCGAGAAAT 766
QY 3362 tacaagtgggcagtgatgggttttagtaacaagaacatttgggcagaggtgggtttggga 3421
DB 767 TGCAGAGTCGCAACGATACCTTTAGT-----ACCATACCTGGAAAGGAGTGGTA 820
QY 3422 aaggtctacaagagcgtctggcagcagcgaactctgtgtctgtcaagaagctcaagaag 3481
DB 821 AGGTGCTTAAGAGGACGCTTGCATGCGTCACCTTGAAGCATTAAGAGCTTAAGAGAG 880
QY 3482 agcgaactccagtgtagagctccagtttcaacaagaagtagagatagatagtcagcag 3541
DB 881 AACGAACACCAAGTGTGAGCTGCAAGTTTCMAACAGAGGTGGAATGATTAGCATGGCTG 940
QY 3542 ttctcgaaactgttagattagaggtttctgtatgaacggaacggaagattgttg 3601
DB 941 TGCATCGAAATCTTCTGCGTCTAGCGTGGTTTCTGCATGACACCAACAGAGCGGCTTGTG 1000
QY 3602 tgtatcttaccatgtgccaatggaagtgtgtgtctgtctcgaagtgtaaaactaaacaa 3661
DB 1001 TATATCCATACATGCTAATGGAAGTGTGCGTGTGTTAAG----- 1044
QY 3662 ttaaacatctgtgtctctctcaattacttgacgtgaagtglttttcatgtttctct 3721

DB 1044 ----- 1044
QY 3722 ttatggttataattgtgtttactaataatgacaagaagagccacgttcaaacctcc 3781
DB 1044 -----AGAGCGTCAGCCATCAAGAACCTCC 1067
QY 3782 gcttgatggccaacgcggaagaagaatcgcgctagctcagctcaggtgttgcttaact 3841
DB 1068 CTTGATTTGGCCAACTGGAAGAGGATTCGATGAGATCTGCTAGGGGGCTTCTTATT 1127
QY 3842 acatgcatcgtcgatccgaagatcattcacgcgtgaacgtaaaagcagaacatccctct 3901
DB 1128 GCATGACCATTTGATTCACAGATATCCATCGATGTAAGAGCCGCAATATATATT 1187
QY 3902 agacgaagaattcgaagcgtgtgttagatctcgggtgtggcgaacgtaactata 3961
DB 1188 GGACGAAGAATTTGAGCGCTGTGTAGCTGATTTGGGTATGCTAGGCTCATGTATACAA 1247
QY 3962 agacatcagtgacacaagcagcagtcggtgcacatcggtcacatcgctccagaatatt 4021
DB 1248 GGATPACCATTTACAACTGCTGAAGGGTACCTTGCGCTACATAGCTCCGAGTACCT 1307
QY 4022 ctcaacgcgaataatcttcagagaagaacgcagcgttttcgatacgaatcgtctaga 4081
DB 1308 CTCGACTGGAAGATCAGACAGAGAGCGATGCTTTGTTAGGATATGCTCTTACA 1367

RESULT 5
T62124
ID T62124 standard; cDNA to mRNA; 3176 BP.
AC T62124;
DR 10-JUN-1997 (first entry)
DE Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
KW Plant; morphogenesis; regulation; short; stem; alteration;
KW Inflorescence; extraneous; gene; expression; transformation;
KW Increase; control; form; length; ds.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT cds 51..2981
FT /tag= a
FT /note= "plant morphogenesis regulatory protein"
PN J09056382-A.
PD 04-MAR-1997.
PE 24-AUG-1995; 216187.
PR 24-AUG-1995; JP-216187.
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
DR MPI: 97-206629/19.
DR P-FSDB; W13408.
PT DNA encoding plant morphogenesis regulatory protein - useful to
PT yield plants with short stems or altered inflorescence
PS Claim 1; Pages 6-10; 17pp; Japanese.
CC The present sequence encodes an Arabidopsis thaliana plant
CC morphogenesis regulatory protein (MRP), which can be used to yield
CC a plant with, e.g. short stems or altered inflorescence. The MRP
CC acts on a plant at a specific site for a specific period, and can
CC therefore be used to regulate extraneous gene expression in a
CC plant. The MRP's cDNA or genomic DNA can be used to transform a
CC plant to increase its MRP expression, and therefore control the
CC form (particularly stem length) of the plant.
SQ Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match 2.6%; Score 108; DB 1; Length 3176;
Best Local Similarity 55.6%; Pred. No. 3.8e-15;
Matches 207; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 3710 tcatgtttccttatttggttataatattgtgtttactaataatgacaagaagagccacc 3769
DB 2201 TCTTCTGTTCTATGACTATTGGAATGTAAGCTCTGGGATCTTCTTCAATGGCCCTAC 2260
QY 3770 gtcaacaacctcgctgattggccaacgcggaagaagaatcgcgctaggctcagag 3829


```
Db 2261 GAAGAAAAAGACTCTGATTGGGACACGCGCTTAAGATAGCATATGGTGCACGACAAAGG 2320
OY 3830 ttgttctacccatcgtacgtacgtccgaagaatccacgtgacgtaaaagcagc 3889
Db 2321 TTTAGCTTATCTACACAGTACTGTAGTCCAGATCATCTACAGAGACGTGAAGTCTCTC 2380
OY 3890 aaacatccctcctagacgaagaattcgaaacggttcttggaagattccggttgccaaagct 3949
Db 2381 CAACATTCCTCTGGACAAAGACTTAGAGGCTGTTGACAGATTTTGGAAATAGCGAAAG 2440
OY 3950 aatgagctataaagacacgtcagtcgaacagcagtcgcgtgacacatcggtcacatcgc 4009
Db 2441 CTTGTGTGTGTAAAGTCACTACATCTACGTTACGTTAGGCGCAGATAGCTTACATAGA 2500
OY 4010 tccagaatctctcaccggaanaatctcagagaanaaacgcagcttttcgatatcggaat 4069
Db 2501 CCCCAGATATGCTCCGACTTCAAGGCTCAGTGAAGAAATCCGATGCTACAGTTATGAAT 2560
OY 4070 catgctctaga 4081
Db 2561 AGTCCCTTTTGA 2572
```

RESULT 6

```
ID T62125 standard; DNA; 9295 BP.
AC T62125;
DT 10-JUN-1997 (first entry)
DE Arabidopsis thaliana plant morphogenesis regulatory protein DNA.
KW Plant; morphogenesis; regulation; short; stem; alteration;
   Inflorescence; extraneous; gene; expression; transformation;
   Increase; control; form; length; ds.
OS Arabidopsis thaliana.
FH Key
FT exon Location/Qualifiers
FT 1803..1881
FT 1882..2227
FT 2228..2366
FT 2367..2467
FT 2468..2539
FT 2540..2643
FT 2644..2715
FT 2716..2809
FT 2810..2878
FT 2879..2968
FT 2969..3040
FT 3041..3118
FT 3119..3190
FT 3191..3266
FT 3267..3338
FT 3339..3421
FT 3422..3493
FT 3494..3586
FT 3587..3655
FT /tag= s
```

```
FT intron 3656..3740
FT /tag= t
FT exon 3741..3812
FT /tag= u
FT intron 3813..3888
FT /tag= v
FT exon 3889..3960
FT /tag= w
FT intron 3961..4048
FT /tag= x
FT exon 4049..4120
FT /tag= y
FT intron 4121..4209
FT /tag= z
FT exon 4210..4281
FT /tag= aa
FT intron 4282..4349
FT /tag= ab
FT exon 4350..4421
FT /tag= ac
FT intron 4422..4508
FT /tag= ad
FT exon 4509..4580
FT /tag= ae
FT intron 4581..4706
FT /tag= af
FT exon 4707..4778
FT /tag= ag
FT intron 4779..4860
FT /tag= ah
FT exon 4861..4932
FT /tag= ai
FT intron 4933..5018
FT /tag= aj
FT exon 5019..5090
FT /tag= ak
FT intron 5091..5176
FT /tag= al
FT exon 5177..5248
FT /tag= am
FT intron 5249..5412
FT /tag= an
FT exon 5413..5481
FT /tag= ao
FT intron 5482..5576
FT /tag= ap
FT exon 5577..5648
FT /tag= aq
FT intron 5649..5726
FT /tag= ar
FT exon 5727..5800
FT /tag= as
FT intron 5801..5882
FT /tag= at
FT exon 5883..6011
FT /tag= au
FT intron 6096..6443
FT /tag= av
FT intron 6012..6095
FT /tag= aw
FT intron 6444..6519
FT /tag= ax
FT exon 6520..6512
FT /tag= ay
FT intron 6891..6974
FT /tag= az
FT exon 6975..7328
FT /tag= ba
FT J09056382-A.
FT PD 04-MAR-1997.
FT PF 24-AUG-1995; 216187.
FT PR 24-AUG-1995; JP-216187.
FT PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
```

PA (CHK-) ZH CHIRYU KANKYO SANGYO GIITSU KENKYU.
 DR WPI; 97-206629/19.
 CC DNA encoding plant morphogenesis regulatory protein - useful to
 PT yield plants with short stems or altered inflorescence
 PS Claim 6; Pages 12-15; 17pp; Japanese.
 CC The present sequence encodes an Arabidopsis thaliana plant
 CC morphogenesis regulatory protein (MRP), which can be used to yield
 CC a plant with, e.g. short stems or altered inflorescence. The MRP
 CC acts on a plant at a specific site for a specific period, and can
 CC therefore be used to regulate extraneous gene expression in a
 CC plant. The MRP's cDNA or genomic DNA can be used to transform a
 CC plant to increase its MRP expression, and therefore control the
 CC form (particularly stem length) of the plant.
 SQ Sequence 9295 BP; 2873 A; 1581 C; 1602 G; 3239 T;

Query Match 2.6%; Score 106.2; DB 1; Length 9295;
 Best Local Similarity 58.3%; Pred. No. 1.4e-14;
 Matches 186; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 3763 ggcacacgtacaacactcgcgttgatggccaacgcggaagaatcgctgctcag 3822
 DB 6520 GCCCTACGAAGAAAGAAAGACTTGATGGACACACGCGCTTAAGATAGCATATGATGACAG 6579
 QY 3823 ctggaagtttgcttacctacatgactgcgtcgcgaagaatcattcaccgtgaagctaa 3882
 DB 6580 CACAAGGTTTAACTTATCTTACACCATGACTGTAGTCCAAAGATCATTACAGAGACGTGA 6639
 QY 3883 aagcagaagaacatcctcttgaacgaagaatcgaagcggttgcttgagattcgggtttg 3942
 DB 6640 AGTGCTCCAACTCTCTTGACAAAGACATTAGAGCGCTGTTGACAGATTGGAATAG 6699
 QY 3943 caaagcgaatggaataaagaacactcagctgacaacagacgctcgtgacatcgctc 4002
 DB 6700 CGAAAGCTTGCTGTGTCAAAAGTCACTTCACTTACGTGATGGGACATGATGGTT 6759
 QY 4003 acatcgctccagaatcctctcacaacgcgaagaatcttcagagaagaacgcgaagcttcg 4062
 DB 6760 ACAATAGACCCGAGATATGCTCGCACTTCACGGCTCATGAGAAATCCGATGCTACAGTT 6819
 QY 4063 acggaatcagctctcaga 4081
 DB 6820 ATGGAATAGTCTCTTGA 6838

RESULT 7
 X07356
 ID X07356 standard; DNA; 4104 BP.
 AC X07356;
 DT 21-MAY-1999 (first entry)
 DE Arabidopsis steroid receptor Bin1 DNA.
 KW BIN1; steroid receptor; receptor kinase; transgenic plant;
 KW brassinosteroid; disease resistance; crop protection;
 OS Arabidopsis thaliana.
 FH Key Location/Qualifiers
 FT CDS 97..3687
 FT CDS /*tag= a
 FN W09859039-A1.
 PD 30-DEC-1998.
 PE 24-JUN-1998; U13100.
 PR 24-JUN-1997; US-881706.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Chory J, Li J;
 DR WPI; 99-081275/07.
 DR P-PSDI; W97819.
 PT New receptor kinase Bin1 involved in brassinolide signalling -
 PT useful for promoting increased yield and disease resistance in
 PT plants and for modulating oocyte maturation
 PS Claim 8; Page 49-52; 72pp; English.
 CC This DNA sequence codes for a novel plant steroid receptor kinase,
 CC designated Bin1 (see W97819), which is involved in the pathway for
 CC the synthesis of the plant steroid hormone, brassinolide. 18 New

CC Arabidopsis dwarf mutants were identified that lacked the ability
 CC to respond to brassinolide, and were named bin mutants. The bin1
 CC mutations were used to map the gene to a small interval on
 CC Arabidopsis chromosome 4. Bin1 was cloned by map-based cloning.
 CC The Bin1 polynucleotide was identified within this interval by
 CC sequencing the wild-type and mutant alleles of this nucleic acid.
 CC Overexpression of Bin1 in transgenic plants provides plants
 CC characterised as having enhanced disease resistance. Increased
 CC plant yield or vegetative biomass and increased seed yield.
 CC Expression of Bin1 may also increase resistance to pesticides.
 CC Inhibition of Bin1, e.g. using antisense oligonucleotides (AON), is
 CC used to render plants male-sterile, and to reduce their stature or
 CC yield, e.g. for creating dwarf varieties. Since Bin1 homologues
 CC may be involved in regulation of the menstrual cycle and uterine
 CC function, Bin1, antibodies and AON may be useful as contraceptives,
 CC for improving success of in vitro fertilisation and to prevent
 CC premature labour. Transgenic animals are also provided, and are
 CC models for studying steroid-receptor interactions or can be used
 CC to screen for therapeutic agents.
 SQ Sequence 4104 BP; 1061 A; 870 C; 930 G; 1243 T;

Query Match 2.1%; Score 87.6; DB 1; Length 4104;
 Best Local Similarity 57.9%; Pred. No. 1.6e-10;
 Matches 175; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 3763 ctgtgttgccaacgcggaagaagaatcgcgctgaagctcagctgagttgcttactcta 3842
 DB 3028 CTTAAATGTCACACGCGGGAAGATTGCGATGATCAGCTAAGGCGCTTGCTTCTT 3087
 QY 3843 catgtaactcgatcgaagaatcattcaccgtgaagtaagaagaacaatcctctta 3902
 DB 3088 CACCAACACTGAGTCCCATATCATCATCAGAGACATGAAATGCCAATATGTTGCTT 3147
 QY 3903 gacgaagaattcgaagcggttgctggaagattcgggttgcaagactaagactataaa 3962
 DB 3148 GATGCAATTTTGGAAGCTCGCGGTTTCAAGATTTTGCAATGGCGAGCTGATGATG 3207
 QY 3963 gacatcagcga---caacagcagtcgctggaacacatcagctcagctcagaatat 4019
 DB 3208 GATACGCAATTTAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 3267
 QY 4020 ctctcacaacgaatcctcagagaagaacgcgaagcttcggaatacgaatcctctta 4079
 DB 3268 TACCAAGTTTCAGGTGTTCAACAAAGACAGCGTTTATGTTACGGTGTGCTTATCTC 3327
 QY 4080 ga 4081
 DB 3328 GA 3329

RESULT 8
 X23531
 ID X23531 standard; cDNA; 3045 BP.
 AC X23531;
 DT 17-JUN-1999 (first entry)
 DE Maize Xa21 gene DT4 cDNA fragment.
 KW Xa21; receptor kinase-like protein; multigene family; RRK; rice; DT4;
 KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
 OS Zea mays.
 PN WO9909151-A2.
 PD 25-FEB-1999.
 PE 17-JUL-1998; U14841.
 PR 13-AUG-1997; US-910386.
 PA (REGC) UNIV CALIFORNIA.
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,
 PI Wang G;
 DR WPI; 99-204431/17.
 PT New RRK polynucleotides and nucleic acid constructs - used for
 PT generating transgenic plants resistant to Xanthomonas
 PS Claim 12; Page 58-59; 67pp; English.
 CC This invention describes a method for conferring disease resistance in
 CC plants. The invention describes the use of novel genes and proteins

QY	1389	cagcttcaacgcggttcccaattactcttclogatctctttttatttggaggggaatcgcta	1448
Db	5204	TT	5145
QY	1449	tctttagtgcgcgatttttattatgaaaattggttggtccttcttgattatgtgaagt	1508
Db	5144	TT	5085
QY	1509	tagcggtagtaacttgaataacacigtcttgccttctctgttcgaataacttgcata	1568
Db	5084	TT	5025
QY	1569	tgtaaagggcatgctcttcttgggttcgaaagcctgggtatttgatatcttaaatgagt	1628
Db	5034	TT	4965
QY	1629	gttgatccaacatctctcigaagaacttacttggttcttcttggattbgaagaattgctt	1688
Db	4964	TT	4905
QY	1689	aattattagcctctaactcagaagagcgctglttgata	1727
Db	4904	TTTTTTTTTTTTTTTGCACGCACTGCAAGGCGTTATGTGATA	4866

RESULT 12
 ID X33184 standard; DNA; 7996 BP.
 AC X33184;
 DT 25-JUN-1999 (first entry)
 DE Base sequence of the plasmid ptx-bcl 2-1-hcd 25.
 KW Copox virus; bcr; viral vector; expression; apoptosis; resistance;
 KW crna; bcl-2; bcl-x1; FLIP; survival; IAP; adenovirus; cancer;
 KW autoimmune disease; graft rejection reaction; inflammation;
 KW inflammatory disease; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN W0991073.A2.
 PD 18-MAR-1999.
 PF 07-SEP-1998; J04010.
 PR 08-SEP-1997; JP-259235.
 PA (RPRG-) RPR GENECENT ASIA PACIFIC INC.
 PI Hamada H;
 DR WPI: 99-243728/20.
 PT New apoptosis-resistant virus-sensitive cell
 PS Example 3; Page 46-49; 51pp; English.
 CC The present invention describes an apoptosis-resistant virus-sensitive
 CC cell line into which an apoptosis resistance gene has been introduced.
 CC The recombinant viruses generated are capable of expressing apoptosis-
 CC associated genes. These can then be used in a variety of diseases for
 CC which the induction of apoptosis is gene transfer, or where the
 CC inhibition of harmful apoptosis is therapeutic. The recombinant viruses
 CC are useful as vectors for gene therapy which can be applied to cancer
 CC therapy for destroying cancer cells selectively, the treatment of
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
 CC encountered the problem where if an adenovirus vector capable of
 CC expressing an apoptosis-associated gene is introduced into animal cells,
 CC the cells producing the virus will be destroyed because the period of
 CC time required to induce cell death by apoptosis is shorter than that
 CC required to replicate and produce the virus, resulting in failure to
 CC obtain a recombinant virus having the integrated apoptosis-associated
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an
 CC apoptosis resistant gene introduced) is established and overcomes the
 CC problem. The present sequence represents the base sequence of the
 CC plasmid ptx-bcl 2-1-hcd 25, which contains the human Bcl-2 gene, and
 CC is used in an example from the present invention.
 SQ Sequence 7996 BP; 2463 A; 2015 C; 1899 G; 1689 T;

Query Match	2.0%	Score 82.2;	DB 1;	Length 7996;
Best Local Similarity	45.5%	Pred. No. 3.4e-09;		
Matches 291; Conservative	0;	Mismatches 348;	Indels 0;	Gaps 0;

QY	1089	ttttatttatttcttttcttcttctgtttcttttaatgctaaggggtttttaaaagggtta	1148
Db	5703	TT	5644
QY	1149	tcsaaaaaatgtagtgagtttggtgtbtagagttgcctcgtgaaagtgtaaatggttgatt	1208
Db	5643	TT	5584
QY	1209	ttcggaagtagtaggttttcgcgactcgaagagcaatcaagattcgaaatttagcat	1268
Db	5583	TT	5524
QY	1269	tgttgattgaaatggagtcgagttatgvgtggttacttactactgactgacttc	1328
Db	5523	TT	5464
QY	1329	cgaaacttcactgctgctgctgcgcgaatttgaaggtcgcggttactcaattact	1388
Db	5463	TT	5404
QY	1389	caagcttaactcgttctcctaattactctcgactcttttacttggaggtgaacgcta	1448
Db	5403	TT	5344
QY	1449	tcttagtgctgcgactttgatttgaataatggtgtgtcttcttgatttgaagatt	1508
Db	5343	TT	5284
QY	1509	tagtggctagactttgaataacacgttttgccttctcgtccagatcaacttctaat	1568
Db	5283	TT	5224
QY	1569	tgtaaagcatggttccttgggttgaanaagctgggttatttgatactaaagattgat	1628
Db	5223	TT	5164
QY	1629	gttgatccaaacactcctcgaagaactcactgttttggtttgtaaagatttgtt	1688
Db	5163	TT	5104
QY	1689	aattatagcctcaatctcagaagagcctgtttgata	1727
Db	5103	TTTTTTTTTTTTTTCACACTGCAAGAGGGTATTATGATA	5065

RESULT	13	
ID	X23533	
AC	X23533 standard; DNA; 3842 BP.	
DT	17-JUN-1999 (first entry)	
DE	Tomato Xa1 clone TRK2 DNA fragment.	
KW	Xa1; receptor kinase-like protein; multigene family; RRR; rice; TRK2;	
KW	plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.	
OS	lycopersicon esculentum.	
PN	W09909151-A2.	
PD	25-FEB-1999	
PF	17-JUL-1998; U14841.	
PR	13-AUG-1997; US-910386.	
PA	(RESC) UNIV CALIFORNIA.	
PI	Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,	
PI	Wang G;	
PI	WPI; 99-204431/17.	
DR	P-PEDB; W23533.	
PT	New RRR polynucleotides and nucleic acid constructs - used for	
PT	generating transgenic plants resistant to Xanthomonas	
PS	Claim 14: Page 61-62; 67pp; English.	
CC	This invention describes a method for conferring disease resistance in	
CC	plants. The invention describes the use of novel genes and proteins	
CC	belonging to the Oryza longistaminata and Oryza sativa receptor	
CC	kinase-like protein (RRR) Xa1 multigene family. Such genes from	
CC	cassava, maize and tomato are also described. The genes and proteins can	
CC	be used for enhancing resistance to Xanthomonas in a plant, preferably in	

CC	rice or tomato.			
Sequence	3842 BP;	984 A;	719 C;	894 G;
				1244 T;

Query Match	1.9%	Score 78.8;	DB 1;	Length 3842;
Best Local Similarity	54.0%;	Pred. No. 1.5e-08;		
Matches 161; conservative	0;	Mismatches 137;	Indels 0;	Gaps 0

Qy	3784	ttgttttgccaaacgcggaagaagaatcgcctcgaagccatgcagagtttgccttcctaac	3843
Db	2973	TGGACTGGAGGGTCTTCACAAATATGCTTGGATATACCCGCTACACTTGGTTACCTGC	3032
Qy	3844	atgtactctcgatccgaagatcattccacggttaagcagaacaacccctctag	3903
Db	3093	ATGATCAGTGTGTACCAACGCTGCTTCATCGTATGTGAAGCCGACCAACATTATTATGG	3092
Qy	3904	acgaagaatccgaagcagtggttgcggaatttcggttcgaagaagtaatggacataaag	3963
Db	3093	ATGAGGAGTATAAGCAATTATTATCTGATTTTGGCTTGGCTAGATTACGGGAACCTTGA	3122
Qy	3964	acactcaatgcacacacagcaagtcctcgtgcgacacatcgttaactcgttcgaagaattcct	4023
Db	3153	AGACCCATGCAACTACACGCTGCTGGCGGGAACCTTGGATATGTTGCTCTGAAATTCGA	3212
Qy	4024	caaccggaaatatcttcagagaaaaacccagcgttttcgatacggatacatcagctcttaga	4081
Db	3213	TGACTTGCCGCGTCTCGAAGACAGGCTGTATCTTACAGATTATGGGGTGTGTGCTTGA	3270

RESULT 14

V06587 standard; cDNA to mRNA; 981 BP

DT	03-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK LRR homologous EST clone.		
DI	receptor kinase; apomixis; apomictic; seeds; production; embryos		
KW	plant breeding; leucine-rich repeat; ss.		
KW	Arabidopsis thaliana		
OS	Key	Location/Qualifiers	
PH	104..760		
FT	CD5		
FT	/*tag= a	/note= 'shows high homology to SERK"	
FT			
PN	W09743427-A1.		
PD	20-NOV-1997.		
PR	13-MAY-1997; E02443.		
PR	14-MAY-1996; GB-010044.		
PA	(NOVS) NOVARTIS AG.		
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;		
DR	WPI: 98-086529/08.		
DR	P-ESDS: W47019.		
PT	Production of apomictic seeds - useful in plant breeding		
PT	Claim 28; Pages 75-77;123pp. English.		
PS	The sequence is that of an EST clone showing high homology to		
CC	SERK LRR (leucine-rich repeat) sequences.		
CC	Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;		
SC			

Query Match	1.88;	Score 72.4;	DB 1;	Length 981;
Best Local Similarity	70.38;	Pred. No. 2.5e-07;		
Matches 97;	Conservative	0;	Mismatches 41;	Indels 0;
				Gaps 0

Qy	1794	aggtaagctcttgacacttgaggtggttactcagttgacacaaatgcttgcag	1853
Db	139	AGGATAGCTCTCTACGCTCTTCGCCGAGTTGACAGATCCAGACATGCTCTCCAG	252
Qy	1854	ctggagatctctcgcgtgagtgaaatctcttcacatggtccatgtracgtgcaaaagagaa	1913
Db	253	CTGGATCTCACTCTGTATATCTTTTACTCTGTTCCATGTCACCTGTATACCAAGACA	312
Qy	1914	cagtgcatagaagatgtga	1931
Db	313	CCGCGTCACTCCTGTGGA	330

RESULT	15
V06586	
ID	V06586 standard; CDNA to mRNA; 1106 BP.

DT	03-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK LRR homologous EST clone.		
KW	receptor kinases; apomixis; apomictic; seeds; production; embryos;		
RV	plant breeding; leucine-rich repeat; ss.		
OS	Arabidopsis thaliana.		
FX	Key	location/Qualifiers	
FT	CDS	142..798	
FT		/tag="	
FT		/note="Shows high homology to SERK"	
PN	W09743427-A1.		
PD	20-NOV-1997.		
PE	13-MAY-1997; E02443.		
PR	14-MAY-1996; GB-010044.		
PA	(NOVS) NOVARTIS AG.		
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;		
DR	WPI: 98-086529/08.		
DR	P-PSDS: M47018.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 28; Pages 71-73; 123pp; English		
CC	The sequence is that of an EST clone showing high homology to		
CC	SERK LRR (leucine-rich repeat) sequences.		
SC	Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;		

Query Match	1.8%;	Score 72.4;	DB 1;	Length 1106;
Best Local Similarity	70.3%;	Pred. No. 2.6e-07;		
Matches	97;	Conservative	0;	Mismatches 41; Indels 0; Gaps 0.

QY	1794	aggttgatgcttgcatactttgaggggttactcgaftgataccaaagtctcttgagag	1853
Db	231	AGGAAATGCTCTCTACGGTCTTCCGSGAGTTTGACAGATGCACGACATGTCCTCCAGAG	290
QY	1854	cttggatcctcagcgaagaaatcctgcacatggtttcattatcactgcgaagaagagaa	1913
Db	291	CTGGGATCCACTCTTTGTTAACTCTTCTTACTGTTTCATGTCACTCTGTACCAAGACAA	350
QY	1914	caatgtcacaagaagtgtta	1931
Db	351	CCGGCTCACTCGGTGGGA	368

Search completed: June 23, 2000, 22:52:13
Job time: 40529 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:42:10 ; Search time 274.21 Seconds
(without alignments)
1934.536 Million cell updates/sec

Title: US-09-180-798-20

Perfect score: 4081

Sequence: 1 tctagaaccttgcattcat.....tacggaatcatgcttctaga 4081

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2-6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2-6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2-6/ptodata/2/ina/5C.COMB.seq:*
4: /cgn2-6/ptodata/2/ina/5D.COMB.seq:*
5: /cgn2-6/ptodata/2/ina/6.COMB.seq:*
6: /cgn2-6/ptodata/2/ina/PCUS.COMB.seq:*
7: /cgn2-6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	72.2	1.8	7218	1	US-08-232-463-14 Sequence 14, Appl
2	71.6	1.8	1554	4	US-08-587-680A-24 Sequence 24, Appl
3	64	1.6	5733	3	US-08-473-553A-1 Sequence 1, Appl
4	59.8	1.5	966	1	US-08-447-185-2 Sequence 2, Appl
5	59.8	1.5	2443	1	US-08-447-185-3 Sequence 3, Appl
6	59	1.4	2571	1	US-07-717-331F-9 Sequence 9, Appl
7	59	1.4	2833	1	US-07-717-331F-1 Sequence 1, Appl
8	55.4	1.4	2749	1	US-07-717-331F-4 Sequence 4, Appl
9	54.2	1.3	2749	2	US-08-265-628-1 Sequence 1, Appl
10	53.2	1.3	19124	4	US-08-487-826B-13 Sequence 13, Appl
11	47.8	1.2	240	1	US-08-628-417-6 Sequence 6, Appl
12	46.2	1.1	3921	4	US-08-587-680A-3 Sequence 3, Appl
13	46.2	1.1	3921	4	US-08-587-680A-3 Sequence 3, Appl
14	46.2	1.1	3921	4	US-08-587-680A-3 Sequence 3, Appl
15	46.2	1.1	6256	3	US-08-475-891A-3 Sequence 1, Appl
16	46.2	1.1	6256	4	US-08-567-375-1 Sequence 1, Appl
17	46.2	1.1	6256	4	US-08-587-680A-1 Sequence 1, Appl
18	46.2	1.1	2447	4	US-09-014-969-14 Sequence 14, Appl
19	45.4	1.1	7218	1	US-08-232-463-14 Sequence 14, Appl
20	45.4	1.1	6243	4	US-09-056-075-1 Sequence 1, Appl
21	44.4	1.1	13473	6	PCT-US96-03916-1 Sequence 1, Appl
22	44.4	1.1	18912	6	PCT-US96-03916-1 Sequence 1, Appl
23	43.8	1.1	2621	4	US-08-553-619B-8 Sequence 8, Appl
24	43.2	1.1	1611	1	US-07-820-011A-3 Sequence 3, Appl
25	43.2	1.1	1611	6	PCR-US93-00445-3 Sequence 3, Appl
26	42.8	1.0	19124	4	US-08-487-826B-13 Sequence 13, Appl
27	41	1.0	5852	1	US-07-867-106-2 Sequence 2, Appl

28	40.6	1.0	2208	6	PCT-US95-08493-1	Sequence 1, Appl
29	40.6	1.0	2580	6	PCT-US95-08493-18	Sequence 18, Appl
30	40.6	1.0	2604	6	PCT-US95-08493-20	Sequence 20, Appl
31	40.6	1.0	2869	1	US-08-374-834-2	Sequence 2, Appl
32	40.6	1.0	2869	2	US-08-644-271-2	Sequence 2, Appl
33	40.4	1.0	2216	2	US-08-046-583-12	Sequence 12, Appl
34	40.4	1.0	2216	3	US-08-384-556A-3	Sequence 3, Appl
35	40.4	1.0	2216	6	US-08-331-355A-24	Sequence 24, Appl
36	40.4	1.0	2216	6	PCT-US94-12364-24	Sequence 3, Appl
37	40.4	1.0	2216	6	PCT-US95-07753-3	Sequence 3, Appl
38	40.2	1.0	642	2	US-08-764-100-13	Sequence 13, Appl
39	40.2	1.0	3000	2	US-08-764-100-9	Sequence 9, Appl
40	39.8	1.0	1920	1	US-08-186-222-1	Sequence 1, Appl
41	39.6	1.0	831	4	US-08-567-375-15	Sequence 15, Appl
42	39.4	1.0	3641	1	US-08-030-096-5	Sequence 5, Appl
43	39.4	1.0	5775	2	US-08-306-691B-15	Sequence 15, Appl
44	39.4	1.0	5775	6	PCT-US93-06251-29	Sequence 29, Appl
45	38.6	0.9	1058	1	US-08-238-163-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F15
; US-08-232-463-14

Query Match 1.8%; Score 72.2; DB 1; Length 7218;
Best Local Similarity 7.3%; Pred. No. 1.8e-08;
Matches 32; Conservative 237; Mismatches 170; Indels 0; Gaps 0;

QY 2462 caacagctctcctggtgcaatctctatgtaacgacccaataatactaaccttcaagtgt 2521
DB 1011 CCAATACGCTCACAGAAATTAATTCGACCTTGCTGCAAGTGAAGGAGCTTGCGATYYY 1070
QY 2522 gtgagctcctcaacttaacttcaactatgtaactcctcccaagtgtgagttgtg 2581
DB 1071 YY 1130
QY 2582 agttaatgcaacttaacttgaatgtaacacagagatctatcaataacacactct 2641
DB 1131 YY 1190
QY 2642 gttcagctcctgcaactgctcctcctcaactcctcaacacacagtgatctatgta 2701
DB 1191 YY 1250
QY 2702 tccctcagtaattcagttgtgtcagtgctgaactatcttgaacttcaatt 2761
DB 1251 YY 1310
QY 2762 ccttgagagtttgcttaactatagacctaigtgacctgttacaagtcaccatgtc 2821
DB 1311 YY 1370
QY 2822 ctgagatcccccgtttctcctccacacacatttatacactcccccagtttccacc 2881
DB 1371 YY 1430
QY 2882 cgagtaagcctcctctt 2900
DB 1431 YYYGTACCAATTTCTCT 1449

RESULT 2
US-08-587-680A-24

; Sequence 24, Application US/08587680A

; Patent No. 5977434

; GENERAL INFORMATION:

; APPLICANT: Ronald, Pamela C.

; APPLICANT: Wang, Guo-Liang

; APPLICANT: Song, Wen-Yuang

; APPLICANT: Szabo, Veronique

; TITLE OF INVENTION: Procedures and Materials for Confering

; TITLE OF INVENTION: Disease Resistance in Plants

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESS: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/587,680A

; FILING DATE: 17-JAN-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/373,375

; FILING DATE: 17-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/475,891

; FILING DATE: 07-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/004,645

; FILING DATE: 29-SEP-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/567,375

; FILING DATE: 04-DEC-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Bastian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 023070-05894005

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1554 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA (partial)

US-08-587-680A-24

Query Match 1.8%; Score 71.6; DB 4; Length 1554;
Best Local Similarity 54.6%; Pred. No. 1.2e-08;
Matches 165; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 3783 ctgattggccaagcggaagagatcgagctcagctcagcgagttgtcttacccta 3842
DB 982 CTCCAATGGGAATAGCTCAATTAATGCAATGCAAGCTGCCAAGGCTTCTTAATTG 1041
QY 3843 catgatactgcatccgaagatcattcaccgtgacgtataaagcagcaacatcctcta 3902
DB 1042 CACCAAGATGCTCCCTATGATATATCCACCGCATGTCACACAAATATATGTTG 1101
QY 3903 gacgaagattcgaagcgtgtgttgagatttgggttggaag--ctaatgacct 3959
DB 1102 AACTGTGAATCTGAAGCTCAATGTCAGATTGTGATAGCCAAAGTACTTGTCAAT 1161
QY 3960 aaagacactcagtgacaagaagcagtcgacacatcgatccacagaatat 4019
DB 1162 GGTACTCTGTGATGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1221
QY 4020 ctctcaaccggaataatcttcaagaaacacgagtttgcgtaaggaatctgtctca 4079
DB 1222 GCATACAGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1281
QY 4080 ga 4081
DB 1282 GA 1283

RESULT 3

US-08-473-553A-1

; Sequence 1, Application US/08473553A

; Patent No. 5859338

; GENERAL INFORMATION:

; APPLICANT: Meyerowitz, Elliot M.

; APPLICANT: Clark, Steven E.

; APPLICANT: Williams, Robert W.

; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,

; TITLE OF INVENTION: Transformed Plants, and proteins

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30


```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
; US-08-473-553A-1

Query Match          1.6%; Score 64; DB 3; Length 5733;
Best Local Similarity 53.6%; Pred. No. 1.8e-06;
Matches 133; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 3712 attgttccttccatggttcataatgtgtgtacactaatgacacagagagccaccgt 3771
DB 4733 atctccttctttatgagtcacatgcttaatggagaccttgagagcttttgcatgatgata 4792
QY 3772 cacaaacctcgttgattgagccagcggaagagaatcgcgctaggtcagctcgaagtt 3831
DB 4793 aaggtggtcatcttcamtgaggagacatgagtagccctggagacctgcamaagggct 4852
QY 3832 tgcttaccatcatatcactcgcagatccgaagatcatccacggtgagtaaaagcagcaa 3891
DB 4853 tctgttatctttacacatgattgtttcaccatttattcttgcatagagattttaagtcacata 4912
QY 3892 aatccctcttgaagcgaagattcgaagcggtgtgttgagatttcggttggcgaagctaa 3951
DB 4913 acattctttggacctctgattttgaagcccatgctgctgattttggccttgctaaagttct 4972
QY 3952 tggactcat 3959
DB 4973 tagttgat 4980

RESULT 4
US-08-447-185-2
; Sequence 2, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATHOGE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,078
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman M., Michael L.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1000
; TELEFAX: (716) -263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-447-185-2

Query Match          1.5%; Score 59.8; DB 1; Length 966;
Best Local Similarity 53.4%; Pred. No. 8.1e-06;
Matches 149; Conservative 0; Mismatches 127; Indels 3; Gaps 1;

QY 3347 ttctcttgaggagatcctacagtgagcgagtgatgagtttaagtaacaagaacatttgagca 3406
DB 86 ttctcttactagattttggaggagacactaaatattatgattcacaagttttaattggac 145
QY 3407 gaagtggtttggaaagctacaaaggagacgcttgagcagcgaactctgttctgca 3466
DB 146 atgctgtcttttggaagcttttaaaaggctgtttccgtaigagaccaaagctgcccctga 205
QY 3467 agagacttgaaggaagagcgaaactcgaagtgtgagagctccagtttcaaacagaagtagaga 3526
DB 206 aaagcggtacacctgactcctcacaagaagtattgaag---agttcgaaacagaaattggaga 262
QY 3527 tgataagtgatgagtcagtcgaaacactgttgagttacagagtttctgtagcacccga 3586
DB 263 ctcttcattttggacagacatccgcatctggtttcattgattagagattcgtgatgaagaaa 322
QY 3587 ccgagagattgcttgatccttaccatgatgccaatgaa 3625
DB 323 atgagatgatttcaatttaataatcattgacagaatggga 361

RESULT 5
US-08-447-185-3
; Sequence 3, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATH
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185

```

FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/111,078
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman Mr., Michael L.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1000
TELEFAX: (716)-263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2443 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-447-185-3

Query Match 1.5% Score 59.8; DB 1; Length 2443;
Best Local Similarity 53.4% Pred. No. 1.3e-05;
Matches 149; Conservative 0; Mismatches 127; Indels 3; Gaps 1;

QY 3347 ttctctcgaggagctacaaagtcgagtcgagtttgaatacaagaacatttgagca 3406
DB 149 TTCTCTTGTAGATTGGAGGACCACTAATATTGATTCACAACTTTTATTGAC 208
QY 3407 gaagttgggttgggaaagtcacaaaggagccttgccagacggaacctgtgtctca 3466
DB 209 ATGGGCTTTGGGAGAGGTTTCAAGGCTGTTGCGATGAGCAAGGTGCGCTCGA 268
QY 3467 agagactgaagaaagagcagctccagtcgagagctccagttcacaagaatagaga 3526
DB 269 AAAGCGCTACACCTGAGCTCTCACAAGTATTGAAG---AGTTGAAACGAAATTGAGA 325
QY 3527 tgataagatgagcagtcacgaacctgtgagattacgaggttctgtaagacaccga 3586
DB 326 CTCCTCATTTTGGAGACATCCGCATCTGTTTCATTGATGATGATCTGTGATAAAGAA 385
QY 3587 ccgagagatgctgtgtatccctacatgagccaatggaa 3625
DB 386 ATGAGATGATTCTAATTTATTAATACATGAGAAATGGGA 424

RESULT 6
US-07-717-331F-9

; Sequence 9, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak

REGISTRATION NUMBER: 26,824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2571 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-07-717-331F-9

Query Match 1.4% Score 59; DB 1; Length 2571;
Best Local Similarity 51.8% Pred. No. 2.1e-05;
Matches 159; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 3778 ctcgctgattggccaacgcggaagagaatcgcgctagctcagctcaggttgcctt 3837
DB 1859 CTAGCTAATTTGGAATGAGATTCGACATTACCAATGGTGTCTCGAGGCTTTTAT 1918
QY 3838 acctacatgatcactgcgagtcggaagatcattcacccgtgacgtlaaagcagaacatcc 3897
DB 1919 ATCTTCATCAAGACTCACGGTTTAGGATATCCACAGAGATTGAAAGTAAGTAACATT 1978
QY 3898 tcttgagcagaatctcgaagcgtgtgttgagatttcgggttgcaagctaatgact 3957
DB 1979 TGCTTGACAAATATGATCCCAAGATCTCGAATTTGGGATGGCCAGGATATTGAAA 2038
QY 3958 ataaagcactcactgacaaacagca---gtccgtgacacatcgttacatcgtccag 4014
DB 2039 GGGAGCAAAAGGAGACTAACAATGAAGGTGTGGAACATAGCGCTACATGTGCCCGG 2098
QY 4015 aatatctcaaccggaataatcttcagagaaacacgcagcttctggataggaatcatgc 4074
DB 2099 AATACCAATGTATGGATATCTCGAATAATCAGATGTTTTCAGTTTGGAGTCATAG 2158
QY 4075 tcttca 4081
DB 2159 TTCTTTGA 2165

RESULT 7

US-07-717-331F-1

; Sequence 1, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2833 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-07-717-331F-1

Query Match 1.4%; Score 59; DB 1; Length 2833;
Best Local Similarity 51.8%; Pred. No. 2.2e-05;
Matches 159; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 3778 ctcccttattgccaagcgcgaagagacgcgtctagctcagctcgagttgtctt 3837
DB 1859 cttaagcttaattggaatgagatcgacatttaccatggtgtctcgagggcttttat 1918
QY 3838 acctacatgatactgctgcatccgaagatcatcaccgtgaagcgttaagcagcaacatcc 3897
DB 1919 attcttattcgaacgtcagcgttttaggattatccacagagatttgaaagtaagtaacattt 1978
QY 3898 tcttaagcagaagattcgaaagcgtgtgttgagatttcggtgttgcaagcgaattagact 3957
DB 1979 tgccttgacaaaattatgattcccaaaagattctcgattttggcagtgccagattttgaaa 2038
QY 3958 ataaagacatcagctgacacagaca--gttcgtgacacatcggtacatcgctccag 4014
DB 2039 ggagacgaaacggaatgacacatgagagtggtcggaacatgacgctacatgctccggc 2098
QY 4015 aatatctcccaacgcgaatcttcaagagaaacgcagcgtttcgagtaagcgaatcgtgc 4074
DB 2099 aataccatgtrattgggatttctcgcaaaaatcagatttttgcatttggaatcattag 2158
QY 4075 ttctaga 4081
DB 2159 ttcttga 2165

RESULT 8

US-07-717-331F-4
; Sequence 4, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2749 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-07-717-331F-4

Query Match 1.4%; Score 55.4; DB 1; Length 2749;
Best Local Similarity 53.2%; Pred. No. 0.00017;
Matches 141; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 3394 aacatttgggcagaggttggttttggaagttctacaaggaagcgttgagcaggaact 3453
DB 1612 AACCAAGCGCGAAGAAAGGCGTTTGCTGTTTCAAGGGAAGGTTAGTGAAGCGGCAA 1671
QY 3454 ctgtgtcttcaagagactgaaggaagcgaactcgaagtggagagctcagttcaa 3513
DB 1672 GAATTTGCAAGTGAAGAGACTATCGAAATGTCAGCTCAAGGTACCGATG--AGTTCATG 1728
QY 3514 acagaagtagagatgataagtagcagttcatcgaacacctgttgagattacgaagtttc 3573
DB 1729 AACGAAGTTAGGCTATGCAAAAGCTTCAGCCACAAATATCTTCCGACTTCTGGCTGT 1788
QY 3574 tgatgacacgcagcagagatgctgtgtatccttacaatgcaatggaagtgttgt 3633
DB 1789 tctgtttatgagggcgaagaaatcttaattttacagatttggagatttaaccttcgat 1848
QY 3634 tctgtctcagaggttaaaactaa 3658
DB 1849 tctcatctctttgagaaacagaa 1873

RESULT 9

US-08-265-628-1
; Sequence 1, Application US/08265628
; Patent No. 5821094
; GENERAL INFORMATION:
; APPLICANT: Rothstein, Steven J.
; TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 W. Madison St. Suite 3400
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,628
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,945
; FILING DATE:
; APPLICATION NUMBER: US 07/847,564
; FILING DATE: 03-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Poehoplen Ph.D., Donald J.
; REGISTRATION NUMBER: 32,167
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-707-8889
; TELEFAX: 312-707-9155
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2749 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Brassica napus
STRAIN: oleifera
INDIVIDUAL ISOLATE: W1
POSITION IN GENOME:
CHROMOSOME/SEGMENT: S-locus
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2574
PUBLICATION INFORMATION:
AUTHORS: GORING, STEVEN J.
AUTHORS: ROTHSTEIN, STEVEN J.
TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A
TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A
TITLE: FUNCTIONAL SERINE/THREONINE KINASE
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749
US-08-265-628-1

Query Match 1.3%; Score 54.2; DB 2; Length 2749;
Best Local Similarity 50.8%; Pred. No. 0.0034;
Matches 156; Conservative 0; Mismatches 148; Indels 3; Gaps -1;

QY 3778 ctccgcttattggtcccaagcggaagaatgcgctagcctcgaagttgctt 3837
DB 1862 CTACGTTAAATGGAAGACAGATTTCACATTACCAATGCTGTGCTGAGGACTTTAT 1921
QY 3838 acctaatgcatcagctgcgaatcgaatcattcaccgtgacgttaaaagcaacaacatcc 3897
DB 1922 ATCTTCATCAGACTCAGCGTTTGAAGTATATCCACAGATATGAAAGTAAGTAACTTT 1981
QY 3898 tctttagaagaagaatcgaagcggtgtgagagattcggtgttggaagaagctaagtcag 3957
DB 1982 TGCTTATATAAATATAGACACCAAGATCTCGGATTTTGGGATGGCCAAATCTTTGCAA 2041
QY 3958 ataagaacctcagctgcgaacagca---gtccgtgcacacatcggtcaccatcgctcag 4014
DB 2042 GGAACGAGACTGAAGCTAACACAGAGAGGTGTCGGAACCTTACGCTCAATGCTCCGG 2101
QY 4015 aatactctcaaccggaataatcttcagagaagaacacgacgttttcggatagcgaatcagc 4074
DB 2102 AGTACGCAATGATGGGATCTCTCGAATAAATCAGATGTTTCACTTTGGAGTCATTTG 2161
QY 4075 ttctaga 4081
DB 2162 TTCTTGA 2168

RESULT 10
US-08-487-826B-13/C
Sequence 13, Application US/08487826B
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487.826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ. ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 1.3%; Score 53.2; DB 4; Length 19124;
Best Local Similarity 48.1%; Pred. No. 0.0016;
Matches 180; Conservative 0; Mismatches 193; Indels 1; Gaps 1;

QY 1336 ttccactgtgcttgccttcctcctaattggaaggctcggtgtactcctaattactcagcttt 1395
DB 15784 TTTTCTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15725
QY 1396 actcggttcacaaatcattcttcgattcttcttatttggaaggatgcgtctctttg 1455
DB 15724 ATTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 15665
QY 1456 tgcctgcatcttgattgaataatgtgtgttcttcttgatttgtaagattagtcgc 1515
DB 15664 TTCTCTCTTTTGTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15606
QY 1516 tagtacttgaaatacactgttttgccttcttccttcagatcaacctgtatattgtaag 1575
DB 15605 TAATTTTCTTTTGAATATCTTTTTCATTTTATTTTATTTTATTTTATTTTATTTTATTTA 15546
QY 1576 gcaatgcttccttggttgaaagcgcggtatttgatattctaagatgatgttgatc 1635
DB 15545 ATTTTATTTATTTTATAAATAATTTTCTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 15486
QY 1636 caaacatctctgaaagactcattgtttttggttttgtaagaagatttctaattatc 1695
DB 15485 AAATTTTCTTTTATTTTATTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 15426
QY 1696 agccctaatctca 1709
DB 15425 CTTTCTTAAATATAA 15412

RESULT 11
US-08-628-417-6/C
Sequence 6, Application US/08628417
Patent No. 5627054
GENERAL INFORMATION:
APPLICANT: GILLESPIE, DAVID
TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC
TITLE OF INVENTION: POLYMERASE CHAIN REACTION

```

; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: U.S. ARMY CHEMICAL AND BIOLOGICAL
; ADDRESSEE: DEFENSE COMMAND
; STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)
; CITY: ABERDEEN PROVING GROUND
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,417
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BIFFONI, DLYSSES J
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 398-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: oligodeoxynucleotide
; HYPOTHEICAL: NO
; ANTI-SENSE: YES
;
US-08-628-417-6

Query Match 1.2%; Score 47.8; DB 1; Length 240;
Best Local Similarity 50.2%; Pred. No. 0.004;
Matches 118; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 1401 ttctcaatactcttcgcgtctcttcttcttctgaagtgatcgctatcttaagtctc 1460
DB 239 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 180
QY 1461 gcatttgattgaagaatgctgtctcttcttgaattgaagttgaagctagta 1520
DB 179 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120
QY 1521 cttagaatacctgttcttcttcttctgtcagatacaacttcttatctgtaagcagc 1580
DB 119 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60
QY 1581 tcttcgttgtaaaagcgtgattatgatacttaagtgatgcttcttctg 1635
DB 59 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5

RESULT 12
US-08-567-375-3
; Sequence 3, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor

```

```

; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1..2676, 3520..3918)
; OTHER INFORMATION: /product= "Xa-21"
;
US-08-567-375-3

Query Match 1.1%; Score 46.2; DB 4; Length 3921;
Best Local Similarity 49.4%; Pred. No. 0.041;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2114 acaaaacaatcaagatttgacagaagaagcactgagttacccttctgtaattgaatc 2173
DB 1054 AAATTTACAAACATTGAACCTTGAGGAAATTAACCTTGAGGAGTTCTTCTATTGCTTT 1113
QY 2174 ttttaacaagttctctatttcttcttcaaggagcttcaagtaacaacatactggccg 2233
DB 1114 TCCAACTTTTCCACTTGCCTTACTTCTTCTTCTGACCTTGAAATGAATGACACAGAAAGC 1173
QY 2234 attctgaataccttggaactctgacaacattagtgagtttgatcttacttaaacgc 2293
DB 1174 ATTCCGAAGGATATGCGCAATCTTATGCTTCAACATCTACTCTGCAACCAACAAT 1233
QY 2294 tttccggtctctatctccgagatcttggaagccttcaagctcagctgattctgagta 2353
DB 1234 TTCAGAGGCTCTTCCATCATCTGTTGGCGAGGCTTAAAACTTAGGCAATCTACTCGCC 1293
QY 2354 tac 2356
DB 1294 TAC 1296

RESULT 13
US-08-567-680A-3
; Sequence 3, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:

```

APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-liang
APPLICANT: Song, Wen-yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,680A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645
FILING DATE: 29-SEP-1995
APPLICATION NUMBER: US 08/567,375
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058940US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product="Xa-21"
US-08-587-680A-3

Query Match 1.1%; Score 46.2; DB 4; Length 3921;
Best Local Similarity 49.4%; Pred. No. 0.041;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2114 acaaaacaattcgaattgacagaagaagcactgagcttaacctttgtaattgaatc 2173
DB 1054 AAATTAACAACATGAACTGGAGAAATAATACCTGGGGAGTCTTCCAAATTCGTTT 1113
QY 2174 ttttaacaagttctattttcttacagggagctttacagtaaacatactacggccg 2233
DB 1114 TCCAAATCTTCCACTTCGCTTAGTTTCTTGCACTTGAAATGAATCAACAGGAGC 1173
QY 2234 attctagaatccttggaactgacaactagtgagtttgatccttaacttaacagc 2293
DB 1174 ATTCCGAAGAGATATGGCAATCTTATGTGGCTTACACATCTCTATCTGCAACAACAT 1233
QY 2294 ttctccggtcctattccggaatcattgggaagcttccaagctgagatttcgtgagta 2353

DB 1234 TTCAGAGGCTCTTCATCATCGTGGGAGGCGTTAAACTAGGATCTTACTGCC 1293
QY 2354 tac 2356
DB 1294 TAC 1296

RESULT 14
US-08-475-891A-3
Sequence 3, Application US/08475891A
Patent No. 5859339
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-liang
APPLICANT: Song, Wen-yuang
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,891A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 02370-058910US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5992 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(512..3149, 3993..4393)
OTHER INFORMATION: /product="RRK-B"
OTHER INFORMATION: /note="Xa21 Xanthomonas spp. disease
OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
sativa)"
US-08-475-891A-3

Query Match 1.1%; Score 46.2; DB 3; Length 5992;
Best Local Similarity 49.4%; Pred. No. 0.051;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2114 acaaaacaattcgaattgacagaagaagcactgagcttaacctttgtaattgaatc 2173
DB 1565 AAATTAACAACATGAACTGGAGAAATAATACCTGGGGAGTCTTCCAAATTCGTTT 1624
QY 2174 ttttaacaagttctattttcttacagggagctttacagtaaacatactacggccg 2233
DB 1625 TCCAAATCTTCCACTTCGCTTAGTTTCTTGCACTTGAAATGAATCAACAGGAGC 1684

QY 2234 attcctagtaacttgcgaacttgcgaacttagtgagtttgatcttacttaaacagc 2293
|||||
Db 1665 ATTCGAGAGGATATTGCAATCTTATGCTTACACATCTCTATCTGCAACACAT 1744
QY 2294 ttctccggtcttattccggaactcattggaagctttcaagctgagattctgtagta 2353
|||||
Db 1745 TTCAGAGGGCTCTTCATCATCGTTGGCGACGCTTAAAACTTAGCATTTCTACTCGCC 1804
QY 2354 tac 2356
|||
Db 1805 TAC 1807

RESULT 15
US-08-475-891A-1
; Sequence 1, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
; OTHER INFORMATION: sativa)"
US-08-475-891A-1

Query Match 1.1%; Score 46.2; DB 3; Length 6256;
Best Local Similarity 49.4%; Pred. No. 0.052;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2114 acaaacacacacagattgacagagaagcactgagttactcttggtaattgaatc 2173

Db 2722 AATTTACAAATCATGTGACTTGGGAGAAAATACCTGGGGAGTTCCTCAATCGTT 2781
|||||
QY 2174 ttcttaacaagttcttatttcttcaaggagagctttacagtaacaataactgccc 2233
|
Db 2782 TCCAACTTTCACCTTCGCTTAGTTTCTTGCACTTGATTGAAATAGATCAGGAGC 2841
QY 2234 attcctagtaacttgcgaactcgaacacttagtgagtttgagattcttacttaaacagc 2293
|||||
Db 2842 ATTCGAGAGGATATTGCAATCTTATGCTTACACATCTCTATCTGCAACACAT 2901
QY 2294 ttctccggtcttattccggaactcattggaagctttcaagctgagattctgtagta 2353
|||||
Db 2902 TTCAGAGGGCTCTTCATCATCGTTGGCGACGCTTAAAACTTAGCATTTCTACTCGCC 2961
QY 2354 tac 2356
|||
Db 2962 TAC 2964

Search completed: June 23, 2000, 22:44:24
Job time: 40210 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:03 ; Search time 6198.48 Seconds
(without alignments)
2668.590 Million cell updates/sec

Title: US-09-180-798-20
Perfect score: 4081
Sequence: 1 tctagaaccttgatcat.....tacggaatcatgctctaga 4081

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*

1:	em_est1:*
2:	em_est2:*
3:	em_est3:*
4:	em_est4:*
5:	em_est5:*
6:	em_est6:*
7:	em_est7:*
8:	em_est8:*
9:	em_est9:*
10:	em_est10:*
11:	em_est11:*
12:	em_est12:*
13:	em_est13:*
14:	em_est14:*
15:	em_est15:*
16:	em_est16:*
17:	em_est17:*
18:	em_est18:*
19:	em_est19:*
20:	gb_est1:*
21:	gb_est2:*
22:	gb_est3:*
23:	gb_est4:*
24:	gb_est5:*
25:	gb_est6:*
26:	gb_est7:*
27:	gb_est8:*
28:	gb_est9:*
29:	gb_est10:*
30:	gb_est11:*
31:	gb_est12:*
32:	gb_est13:*
33:	gb_est14:*
34:	gb_est15:*
35:	gb_est16:*
36:	gb_est17:*
37:	gb_est18:*
38:	gb_est19:*
39:	gb_est20:*
40:	gb_est21:*
41:	gb_est22:*
42:	gb_est23:*
43:	gb_est24:*
44:	gb_est25:*

45:	gb_est26:*
46:	gb_est27:*
47:	gb_est28:*
48:	gb_est29:*
49:	gb_est30:*
50:	gb_est31:*
51:	gb_est32:*
52:	em_est20:*
53:	em_est21:*
54:	em_est22:*
55:	em_est23:*
56:	em_est24:*
57:	em_est25:*
58:	em_est26:*
59:	gb_est33:*
60:	gb_est34:*
61:	gb_est35:*
62:	gb_est36:*
63:	gb_est37:*
64:	gb_est38:*
65:	em_est27:*
66:	em_est28:*
67:	em_est29:*
68:	em_est30:*
69:	gb_est39:*
70:	gb_est40:*
71:	gb_est41:*
72:	gb_est42:*
73:	gb_est43:*
74:	gb_est44:*
75:	em_est31:*
76:	em_est32:*
77:	em_est33:*
78:	em_est34:*
79:	gb_est45:*
80:	gb_est46:*
81:	gb_est47:*
82:	gb_gss1:*
83:	gb_gss2:*
84:	gb_gss3:*
85:	gb_gss4:*
86:	em_gss1:*
87:	em_gss2:*
88:	em_gss3:*
89:	em_gss4:*
90:	gb_gss5:*
91:	gb_gss6:*
92:	gb_gss7:*
93:	gb_gss8:*
94:	gb_gss9:*
95:	em_gss5:*
96:	em_gss6:*
97:	em_gss7:*
98:	em_gss8:*
99:	em_gss9:*
100:	em_gss10:*
101:	em_gss11:*
102:	gb_gss10:*
103:	gb_gss11:*
104:	em_gss12:*
105:	gb_gss12:*
106:	gb_gss13:*
107:	gb_gss14:*
108:	gb_gss15:*
109:	gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

```

Arabidopsis thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenomc.bio.upenn.edu
Seq primer: 7'
Class: BAC ends
High quality sequence start: 76
High quality sequence stop: 704.

FEATURES
    source
        1..1202
        /organism="Arabidopsis thaliana"
        /strain="Columbia"
        /db_xref="taxon:3702"
        /clone="T2E10"
        /clone_11b="TAMU"
        /sex="hermaphrodite"
        /note="Vector: BelobAC11; Site_1: HindIII; Site_2:
        HindIII; Produced by Rod Wing"

BASE COUNT      322 a      275 c      238 g      354 t      13 others
ORIGIN

Query Match      5.7%; Score 233.8; DB 84; Length 1202;
Best Local Similarity 66.6%; Pred. No. 4,4e-39;
Matches 428; Conservative 0; Mismatches 146; Indels 69; Gaps

QY 3002 aattgattccaggtggtgatgataactcagacaaagctgtgtgagttgctgcaggt 3061
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 AATTAATTTTAAGTGTGGATAGTCTCTACTGTGAGCCATCGGAGACGATGTGCTGTGT 187

QY 3062 gctgtgtgctctcttgcgtctccctcgaatagcccttgcctgtgtggtgcgacgaagccaa 3121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 GCTGCTTTACTATTTCCTCCCTCCCTTTAGCGCTTTTGCTTGCTGGCGTAAAAAAAACCT 247

QY 3122 ctgagatcttctcgaatgctgcgtgtagtttatctatctgcattagttctcttcttag 3181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 CAAAGAAATCTCTCTTTGATGTTCCTGGTAGTCACCT-----GAGTCTGCATATTC 296

QY 3182 ccagcaatttctgttcgcgaaagatattggaacaaactgttaatgaaataacatacataa 3241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 CAAAGCTTTGTTTCATTTTCAGAAATTCGAAATTCGAAATTC-----A 326

QY 3242 gtcaattgtttttaagattcaaacctcttttgatgataaatctcgattgcaaatctctaa 3301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 2
LOCUS AQ969360 714 bp DNA GSS 28-JAN-2000
DEFINITION LERM53TR LERG Arabidopsis thaliana genomic clone LERM53, genomic survey sequence.
ACCESSION AQ969360
VERSION AQ969360.1 GI:6797061
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 714)
Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T., Feldblyum,T., Liang,F., Greasy,T. and Fraser,C.M.
Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
JOURNAL Unpublished (2000)
COMMENT On Dec 15, 1999 this sequence version replaced gi:4215892.
CONTACT: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ateligr.org
For additional information, see <http://www.tigr.org/tdb/at/at.html>
Seq primer: TR
Class: shotgun.
FEATURES
source Location/Qualifiers
1..714
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERM53"
/note="Organ: leaf; Vector: pUC19UK; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."
BASE COUNT 205 a 177 c 112 g 220 t
ORIGIN
Query Match 5.2%; Score 213; DB 92; Length 714;
Best Local Similarity 60.9%; Pred. No. 1e-34;
Matches 389; Conservative 0; Mismatches 210; Indels 40; Gaps 1;
QY 3443 cagacggaactctgtgtcgtcaagaagactgaagaagacgaactcgaagtggagagc 3502
DB 702 CGATGGCACCACCAAAATCGCTGTAAGATGATTTGAACTCCAGAGGAGATG 643
QY 3503 tccagttcaacaagaagtgaagtatagtatgagcttcacgaactctgaactcttgaagt 3562
DB 642 AAGCTTCCAGAGAGAGTGAAGTGAATGATGCTTCAAGAACTCTCTCGCC 583
QY 3563 taagaagttctgtatgacacgacgaagagattgtctgtatccctacatgagccaagc 3622
DB 582 TTATCGGCTTTGTGACAAACAACTGAAGACTTTGTGATCTTTCATTCAGAAATC 523
QY 3623 gaagtgttgcgtcgtcctcaagagtaaaaactaaacaattaaacattgtgctctc 3682
DB 522 TAACTGTTGATATTTGCTTAAGAGGTAAAAAAGAACATNA----- 480
QY 3683 tcaattacttgaagtgaagtgttttctgtttcccttaaggttcaataattgttg 3742
DB 480 -----GCTATTAACCTTTTGTATTTCAAGAGATTAAAGATTTTGTG 443
QY 3743 ttacactaatgacacagagagggccacccgtcaccaactcgccttgattggccaagcgagaa 3802
DB 442 ATCTATATCTATCTGACAGAGATTAAACCCGGGATCCAGTTCTGATTTGTTCAAGAGAA 383

QY 3803 gagaatcgctagctgcctcgcctcgcctcgcctcgttcttactacatgatactgcgtccgaa 3862
DB 382 ACAGATTGCTGTAGTGCAGACGAGACTCGAATATCTTCAATGAACATTCGAACCCAA 323
QY 3863 gatcatcaccgtgaagtaaaacacacacatcccttaagcaggaattcgaagcgt 3922
DB 322 GATCATACAGAGATGTGAAGCTCAAAATGTTTACTAGATGAAGACTTTGAAGAGT 263
QY 3923 tgttgagattcgtgggttgcaagactaaatgaactataaagaactcagtgacacagc 3982
DB 262 GGTGGGATTTGTTGTTAGCCAAAGTTGGTAGATGTTGAAGACTATGTAAACCACTCA 203
QY 3983 agtccgtgaccatcggtacacatgctccagaatactcctcaacggaaacttcaga 4042
DB 202 GTTCCAGAGAACATGGGTCATATTTCCACAGATGTATATTCACAGGGAATTCGTGGA 143
QY 4043 gaaacgcagcttttcgagatcgaatcatgcttctaga 4081
DB 142 GAATACCGATGTTTGGGATCGAGATATGCTTGTGA 104
RESULT 3
LOCUS AQ969361 621 bp DNA GSS 28-JAN-2000
DEFINITION LERM53TR LERG Arabidopsis thaliana genomic clone LERM53, genomic survey sequence.
ACCESSION AQ969361
VERSION AQ969361.1 GI:6797062
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 621)
Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T., Feldblyum,T., Liang,F., Greasy,T. and Fraser,C.M.
Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
JOURNAL Unpublished (2000)
COMMENT On Dec 15, 1999 this sequence version replaced gi:4215893.
CONTACT: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ateligr.org
For additional information, see <http://www.tigr.org/tdb/at/at.html>
Seq primer: TR
Class: shotgun.
FEATURES
source Location/Qualifiers
1..621
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERM53"
/note="Organ: leaf; Vector: pUC19UK; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."
BASE COUNT 191 a 102 c 154 g 173 t
ORIGIN
Query Match 5.2%; Score 210.4; DB 92; Length 621;
Best Local Similarity 60.6%; Pred. No. 3.5e-34;
Matches 387; Conservative 0; Mismatches 212; Indels 40; Gaps 1;
QY 3443 cagacggaactctgtgtcgtcaagaagactgaagaagacgaactcgaagtggagagc 3502
DB 21 CGATGGCACCACCAAAATCGCTGTAAGATGATTTGAACTCCAGAGGAGATG 80

Db	403	TGAMATGCTGATGCTTCTTGA	426
RESULT	5		
LOCUS	A1486453		
DEFINITION	A1486453	471 bp	mRNA
ACCESSION	EST244774	tomato ovary, TAMU Lycopersicon esculentum cDNA clone	EST
VERSION	CLD862	mRNA sequence.	29-JUN-1999
KEYWORDS	A1486453.1	GI:4381824	
SOURCE	EST.		
ORGANISM	tomato.		
REFERENCE	Lycopersicon esculentum		
AUTHORS	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumphylophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.		
TITLE	1 (bases 1 to 471)		
JOURNAL	Alcala, J., Vrebalov, J., White, R., Matern, A.L., Viston, T., Holt, L.E., Liang, F., Opton, J., Konning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.		
COMMENT	Generation of ESTs from tomato carpel tissue Unpublished (1999) On May 7, 1998 this sequence version replaced gi:3121325. Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU. Location/Qualifiers 1..471 /organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081" /clone="CLD862" /clone_id="tomato ovary, TAMU" /tissue_type="carpel" /dev_stage="5 days pre-anthesis to 5 days post-anthesis" /lab_host="X11-blue MRF" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; CLD - Tomato Carpel EST Library. Oligo(dT)-primed and directionally cloned cDNA in vector lambda ZAP II with 5' and 3' ends located at the EcoRI and XhoI sites, respectively"		
BASE COUNT	136 a	88 c	112 g 135 t
ORIGIN			
Query Match	4.2%;	Score 169.4;	DB 47; Length 471;
Best Local Similarity	72.0%;	Pred. NO. 1.4e-25;	
Matches 221;	Conservative	0;	Mismatches 86; Indels 0; Gaps 0;
3775	aactcgcgtgattgagccaacgcggaagagaatcgcgctagagtcacgtcaggtt	3834	
70	AGCGCTTTGTAATGGCCAACTAGAGACGCGGTGGGTGCTAGCTGACCTTG	129	
3835	cttactacatgatcactcgtcgatccgaagatcattcaaccgtgacgtaaagca	3894	
130	AATACCTCATGTAACACGTAAATCCAAAGATTAATTCACCGCTATGTAAAG	189	
3895	tcctcttgagcaagaattcggaagcgtgtgtttgaggaattcggttggaagc	3954	
190	TGTTACTAGATGAATTTTGAAGCTGTAGTTGGTGTATTTTGCCCTGGCAAG	249	
3955	actataagacactcagctgacacagcagcgtcgtgacacatcgttcacgtc	4014	
250	ACGTTAAGAAACCAATGAGCTACTCAAGTTCGTGCTACATGGGCCATATAG	309	
4015	aatactctcaacgcggaatacttcaagagaaacgaagtttcggatcaggaat	4074	

Db	310	AAATCTATTCACATGGCAAAATCATCAGAAAAAATACTGATGTTTGGCTATGAAATCATGC	369
QY	4075	ttctaga 4081	
Db	370	TTTTGCA 376	
RESULT	6		
LOCUS	AA738544	991 bp	mRNA
DEFINITION	SbRLK2 Sorghum bicolor cv. TX430 leaf		EST 17-AUG-1998
ACCESSION	AA738544		
VERSION	AA738544.1	GI:3421469	
KEYWORDS	EST.		
SOURCE	Sorghum.		
ORGANISM	Sorghum bicolor		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.		
ATTNORS	1 (bases 1 to 991)		
TITLE	Annon,F., Chang,J.-I., Paterson,A.H. and Stockhaus,J. Characterization of 14 different putative protein kinase cDNA clones of the C4 plant Sorghum bicolor		
JOURNAL	Mol. Gen. Genet. 259 (1), 115-122 (1998)		
MEDLINE	98409267		
COMMENT	On Nov 29, 1993 this sequence version replaced gi:36080. Contact: Stockhaus J Institut fuer Entwicklungsbiologie und Molekularbiologie der Pflanzen Heinrich-Heine-Universitaet Universitaetsstrasse 1, 40225 Duesseldorf, Germany Tel: (49)-211-81-14953 Fax: (49)-211-81-14871 Email: stockhaeuni-duesseldorf.de		
FEATURES	Insert length: 1500 Std Error: 0.00.		
source	Location/Qualifiers		
	1..991		
	/organism="Sorghum bicolor"		
	/strain="cv. TX430"		
	/db_xref="taxon:4558"		
	/clone="SbRLK2"		
	/clone_lib="Sorghum bicolor cv. TX430 leaf"		
	/dev_stage="green"		
	/note="Organ: leaf; Vector: Lambda ZAP II; Site.1: Eco RI; Site.2: Xho I; leaf cDNA library from green leaves, unidirectionally cloned"		
BASE COUNT	253 a 194 c 293 g 251 t		
ORIGIN			
Query Match	4.1%; Score 167; DB 38; Length 991;		
Best Local Similarity	55.7%; Pred. No. 4.0e-25;		
Matches 428; Conservative 0; Mismatches 225; Indels 116; Gaps 1;			
Db	3313	gattccagaagttcatctcttggagacgtcaagggtttcttcttggggagagctcaagtgccg 3372	
QY	3373	agttcaggttttagtacaagaacattttgggcaaggttggtttgggaagctctacaag 3432	
Db	61	ACGATGATATTTCGTCGTAGAAAAATGTTCTTGGACAAAGGGGCTTTGGGAAAGTATTTAA 120	
QY	3433	ggaagcttggcagaaggaactctgttgcgttcaagaagactgaaggaagagcgaactcca 3492	
Db	121	GGAGCACTTCCTCCAGATGGCACTAGATTCTCTGTAAACGGTTAACTGATTACGAAATCCT 180	
QY	3493	gtgtggagagctccagtttcaaacagaagtaagatgataagatagtcgaagcttcgaaac 3552	
Db	181	GGTGGAGAGGCTGCTTTCTTGGTGAAGGTTAGAGTGATATGATGTGCAAGTTACACCGAAT 240	

```

OY 3553 ctgtgagataacagagttctgtatgacacgaccagagagattgctgtatccttcac 3612
    || || || || || || || || || || || || || || || || || || || || ||
Db 241 CTTTAAAGATTGATTGGTTCTTCTACACACAAACAGAGCGCTGCTTTTATCTTTC 300
OY 3613 atgccaatggaagtgtgtctgtctgtctcagaagaaataaacaataaacatctt 3672
    || || || || || || || || || || || || || || || || || || || || ||
Db 301 ATGCACAAATCTTAGTGT..... 319
OY 3673 gtgtctctctcaataacttgacgtgaagtgtttttcatgttttccattatggttca 3732
    || || || || || || || || || || || || || || || || || || || || ||
Db 319 ..... 319
OY 3733 taattgtgttacactaatgacacagagagcaccgacacacccctcgttgatggc 3792
    || || || || || || || || || || || || || || || || || || || || ||
Db 319 -----TCTACCGTCTAGAGAAATTAACTCGGAGCCAAATTAGATTGCT 364
OY 3793 caacgcggaagaagaatcgcagctcagctcagctcagctcagctcagctcagctc 3852
    || || || || || || || || || || || || || || || || || || || || ||
Db 365 CTGCAAGGAGAGCGAGTGGCTATAGGCAACAGCTGAGATGTTGACAGAGCACT 424
OY 3853 ggcagtcggaagatcattccaccgtagcgtgaaagcaaacacatcccttaagcgaagaat 3912
    || || || || || || || || || || || || || || || || || || || || ||
Db 425 GCATCTTAAGATTATACATCGTACGTCAGAGGCTGCAATGCTTGTGATGAAGGTT 484
OY 3913 tcgaagcgtgtgtgagagattcgggtgtgcaagcgaagcgaagcgaagcgaagcga 3972
    || || || || || || || || || || || || || || || || || || || || ||
Db 485 TTGAACCGGTTGTGGATTGGGCTTGCGCCAGCTGCTGATGATGACAGAGACTCG 544
OY 3973 tgcgaacagcagtcggtgacacatcggttcacatcgctccagaatactctcaacgcga 4032
    || || || || || || || || || || || || || || || || || || || || ||
Db 545 TGACTACTACGTCCTCGTGAAACATAGGTCACATGCCCCGTGAAGATTGTCACAGGGA 604
OY 4033 aacttcagaagaacacgagcttttcgatacggaatcgccttcaga 4081
    || || || || || || || || || || || || || || || || || || || || ||
Db 605 AGTCATCCGAGAACCGAGTGTGTTGCTTACGCGCAATGCTTCTCGA 653

RESULT 7
AM329855 .475 bp mRNA EST 28-JAN-2000
LOCUS N201132e rootphos(-) Medicago truncatula cDNA clone MHRP-24H06,
DEFINITION mRNA sequence.
ACCESSION AM329855.1 GI:6800450
VERSION AM329855.1
KEYWORDS EST.
SOURCE Medicago truncatula
ORGANISM Medicago truncatula
REFERENCE 1 (bases 1 to 475)
AUTHORS Harrison,M.J., Liu,J., Gonzales,M.B. and Ellis,L.
TITLES ESTs from phosphate starved roots
JOURNAL Unpublished (1999)
COMMENT On May 1, 1997 this sequence version replaced gi:2059389.
Contact: Maria J. Harrison
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401, USA
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
Developmental stage: phosphate starved; Date: 01/23/00; Submitted
to the database of Expressed Sequence Tags (dbEST) on 01/26/00;
More information is available at
'http://chrystle.tamu.edu/medicago'.
Seq primer: T3.
FEATURES
Source location/Qualifiers
1..475
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"

```

```

/clone="MHRP-24H06"
/clone_lib="rootphos(-)"
/tissue_type="roots"
/Note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The
cDNA was directionally ligated into the Unizap XR vector
from Stragene and packaged using Gigapack IIR Gold
packaging extracts. Plasmids containing cDNA inserts
were excised from the recombinant lambda-zap phage using
Ex-assist helper phage and propagated in SOUL cells."
BASE COUNT 170 a 74 c 107 g 124 t
ORIGIN

```

```

Query Match 4.1%; Score 166.2; DB 80; Length 475;
Best Local Similarity 71.3%; Pred. No. 6.7e-25;
Matches 219; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

```

OY 3775 aactcgcgttgatggcgaacgcggaagaatcgcgtcgaagctcagctcagctgtgt 3834
    || || || || || || || || || || || || || || || || || || || || ||
Db 68 AGCAGGCTTTAGACTGGGCAACAGAGAGAAATGACATTAGAGCAGAGAGAGGTTAC 127
OY 3835 ctaccctacatgacactgcagtcgaagatcattcaccgtgacgtgaaagcagaaca 3894
    || || || || || || || || || || || || || || || || || || || || ||
Db 128 TATATTACATGAACAGTGTGATCCAAAGATATTCACAGAGATGTTAAAGCAGCAATA 187
OY 3895 tcccttaagcgaagaattcgaagcgtgtgttggaattcgggttggaagcgaatag 3954
    || || || || || || || || || || || || || || || || || || || || ||
Db 188 TATTGCTAGATGATTATGAGGCTGTACTAGACACTTGGATTAGCAAACTATTGG 247
OY 3955 actataagaacatcactgacacagcagctcgtgacacatcgcgttcacatcgcag 4014
    || || || || || || || || || || || || || || || || || || || || ||
Db 248 ACCAGAGAGACTCATGATGACAAACAGCACTGAGAGAACTGATGACACATAGCCCCG 307
OY 4015 aatatctcaaacgcgaatacttcagagaagaacgcgatttcggtatcgaatcagtc 4074
    || || || || || || || || || || || || || || || || || || || || ||
Db 308 AATACCTATCAACAGCAATCCTCAGAGAAACAGATGATTGGATTGGAAATCTTC 367
OY 4075 ttctaga 4081
    || || ||
Db 368 TGCTTGA 374

```

```

RESULT 8
AA738547 656 bp mRNA EST 17-AUG-1998
LOCUS SBRK5 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
DEFINITION SBRK5 similar to protein kinase, highest similarity to
receptor-like protein kinases, mRNA sequence.
ACCESSION AA738547
VERSION AA738547.1 GI:3421472
KEYWORDS EST.
SOURCE Sorghum.
ORGANISM Sorghum bicolor
REFERENCE 1 (bases 1 to 656)
AUTHORS Anen,F., Chang,J.-L., Paterson,A.H. and Stockhaus,J.
TITLES Characterization of 14 different putative protein kinase cDNA
clones of the C4 plant Sorghum bicolor
JOURNAL Mol. Gen. Genet. 259 (1), 115-122 (1998)
MEDLINE 98409267
COMMENT On Nov 29, 1993 this sequence version replaced gi:636097.
Contact: Stockhaus J
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49)-211-81-14953
Fax: (49)-211-81-14871
Email: stockhaeuni-duessel.dorf.de

```

Insert Length: 1100 Std Error: 0.00.

FEATURES	location/Qualifiers
source	1. .656

```

/organism="Scorghum bicolor"
/strain="cv. TX430"
/db_xref="taxon:4558"
/clone="SbRLK5"
/clone_lib="Scorghum bicolor cv. TX430 leaf"
/dev_stage="green"
/note="Organ: leaf; Vector: lambda ZAP II; Site:1: Eco RI; Site:2: Xho I; leaf cDNA library from green leaves, unidirectionally cloned"

```

BASE COUNT	161 a	145 c	184 g	166 t
ORIGIN				

Query Match	4.13;	Score 165.8;	DB 38;	Length 656;
Best Local Similarity	73.48;	Pred. NO. 8.1e-25;		
Matches 212;	Conservative 0;	Mismatches 77;	Indels 0;	Gaps 0;

QY	3793	caagcggaagaagatctggcgtcgtgagccacgtcgtaggttgccttaccacagtgtaact	3855
Db	122	CAACAGAGAGACGGATGTGCGCTTGATGTCCAAAGAGACTTCTTCTTGCTGATGACT	181
QY	3853	gcgatacgaagaatcattaccgtgtacgttaaagcagcaaacalcctcttagacgaagaat	3912
Db	182	GGATCCCAAAATATCATTCATGGGATCTCAAGCTGCAAAATATCTTTCATGAGGACT	241
QY	3913	tcgaaagcggtgttgtagagatttcgggtgttgccaagcctaagcgcttaagaacatcagc	3972
Db	242	TCGAGGACAGTTGTGGGTATTTTGGGCTTGCCAAAGCTTATGAGCTACAAAGATACCAGT	301
QY	3973	tgacacacagcagtcctgtgacccatcgtgtacatcgtctccagatactctcaaacggaa	4031
Db	302	TCACAACTGCTGCTCGTGGAGCAATGAGACATTTGGCCTGAGTACTATCACTGCA	361
QY	4033	aactctcagagaaaaacgagcttttcggatacaggaatcattcgtctctaga	4081
Db	362	ACTCTCTCGAAAGACTGATGTTTTTGGCTATGGAGCAATGCTTTCGGA	410

RESULT	9			
AA738546				
LOCUS	AA738546	655 bp	mrna	EST
DEFINITION	SbRRK4 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone			
	SbRRK4 similar to protein kinase, highest similarity to receptor-like protein kinases, mRNA sequence.			

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 655)
Annen, F., Chang, J.-L., Paterson, A. H. and Stockhaus, J.
Characterization of 14 different putative protein kinase cDNAs
clones of the C4 plant *Sorghum bicolor*
Mol. Gen. genet. 259 (1), 115-122 (1998)

On Jan 13, 1998 this sequence version replaced g1:341414.1.
Contact: stockha@uni-duesseldorf.de
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49)-211-81-14953
Fax: (49)-211-81-14871
Email: stockha@uni-duesseldorf.de

Insert Length: 600 Std Error: 0.00.

FEATURES

Source

```

1..655
/organism="Sorghum bicolor"
/strain="cv. TX430"
/db_xref="taxon:4558"
/clone="SBRUK4"
/clone_1fb="Sorghum bicolor cv. TX430 leaf"
/day_stage="green"
/site="Ogo 1; leaf, Vector: Lambda ZAP II; Site_1: Eco RI
/site_2: Xho I; leaf cDNA library from green leaves,
```

BASE COUNT	169 a	124 c	179 g	183 t
ORIGIN				

Query Match	4.0%;	Score 165.2;	DB 46;	Length 655;
Best Local Similarity	56.1%;	Pred. No. 1.1e-24;		
Matches 416; Conservative	0;	Mismatches 213;	Indels 113;	Gaps 1.

QY	3340	aaggggtttcttggcgsgggtcaacaagtgcgagtgatgaggttttaagtaacaagaacct	3399
Db	4	AAAAATTGATGGATGGCAAGATTGCACCTTGGCACTGTAATAATTGATGAGAAAAATTT	63
QY	3400	ttggcgcaagtggtggttttggnaagtcctacaagggagcgttggcagacgaactctgtt	3459
Db	64	CTTGGACAAAGGGGGCTTTGGGAAAGTATATAAGAGACACTTCCAGATGGCACTAAGTT	123
QY	3460	gctgtcaagagactgtaaaggaagacgaactccaggtgtggaagctccagtttcaaacagaa	3519
Db	124	GCTGTAAACGGTTAACTGATTACGAAGTCTGTGGAGAGGGTCTTTCTTGGCTGAG	183
QY	3520	gtagagatgatgaatgatgtgcagcttcatcogaacctgttagattacagaggttctcgtatg	3579
Db	184	GTTTGAGCTGATTAGTGTGTCGACGTTCACCGGAATCTTTAAGATTGATTTGGTTTGTGTACA	243
QY	3580	aaaccgaccggaggaattgctgtgtatcccttaacatgccaatgggaagtgttgcttcgagt	3639
Db	244	ACACAAACAGAGCGCGCTTGTTGTTATCTTTATCCGAGAACTTATAGTGTGGCTAACCGT	303
QY	3640	ctcgaagtgtaaaacataacaattaaacatcttgctctctctcaattacttgagtg	3699
Db	304	CTACGAGAAATTAA-----	318

QY 3700 aagtggttttcattgtttcccttatggtcataattgtgttaacactaatgacacag 3755
Db 318 ----- 318

QY 3760 agagggccacgcgtcacacctccgcttattggtgcacagcggaaagagaatcgcgctagct 3815
||| ||| | ||||| | | ||||| | ||| |||||
Db 318 -----ACCTGGGGAGCCCAATATTAGATTGGTCTGTGCAGGAACGAGTGGCTATTAGGCA 370

Oy 3820 cagctcgaagtgttgccttaacctacatgatacatctgcgatccgaagaatcattcacccgtgacc 3879

Dh 371 CAGCTCGTGGACGTGAGATATATGCCACGACGACACGCAAACTCCTAAGATTATACATCGTGAGC 430

DY 3880 taaagcagcaacatctcttagcagaagtctgaagcggtgtytgagattcggct 3935
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 3940 tggcaagctaattgactataaaagacactcacgtgacaacagcgtccgtggcacatcg 3998
||||| ||||| ||||| | ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Oy 4000 gtacacatcgcttcagaatatcttcaaccggaatacttcagagaaaaccgcgttttcg 4055
||||||| ||| ||||||| ||| ||| ||| ||| ||| ||| ||| |||

DD	551	GTACCAATGCCCCCGAATATTTGCCC
DY	4060	gatacggaatcatgtcttcta 4081

RESULT 10
D25047

LOCUS	25047	299 bp	mrna	EST	08-JUL-1999
DEFINITION	R25047.6	Rice root	Oryza sativa	cdna	clone R2976_1A, mRNA
ACCESSION	D25047				
VERSION	D25047.1	GI:428891			
KEYWORDS	EST.				
SOURCE	Oryza sativa.				
ORGANISM	Oryza sativa				
REFERENCE					
AUTHORS	Minobe, Y. and Sasaki, T.				
TITLE	Rice cdna from root				
JOURNAL	Unpublished (1995)				
COMMENT	Contact: Yuzo Minobe National Institute of Agrobiological Resources Rice Genome Research Program 2-1-2 Kannondai, Tsukuba Ibaraki, Japan 305 Tel: 0298-38-7441 Fax: 0298-38-7468 Email: minobe@erts0.riken.go.jp PROJECT = "Rep".				
FEATURES					
source	Location/Qualifiers 1..299 /organism="Oryza sativa" /strain="Nipponbare, sub-species japonica" /db_xref="taxon:4530" /clone="R2976_1A" /clone_1lb="Rice root" /note="Prepared from seedling root."				
BASE COUNT	93 a	54 c	67 g	85 t	
ORIGIN					
Query Match	4.0%	Score 164.4	DB 20	Length 299	
Best Local Similarity	74.5%	Pred. No. 1.6e-24			
Matches 207	Conservative 0	Mismatches 71	Indels 0	Gaps 0	
Db 3	AGGATCGCGTAGGCTTCAGCTCAGGTTGTCTTACCTACATGATCAGTCCGATCCGAAG				
Qy 3804	agaatcgcgtaggctcagctcagctcaggtttgtcttacctacatgatactcagtcgacgaag				
Db 63	ATCATCATCGTGTATGTCNAAGCGCGCAATATTTATTGATGAGAGACTTTGAACTGTA				
Qy 3864	atcatcaccgttgacgtaaaagacgaacacatccctcttagacgaagaattcgaaagcgtt				
Db 63	ATCATCATCGTGTATGTCNAAGCGCGCAATATTTATTGATGAGAGACTTTGAACTGTA				
Qy 3924	gttgacgatttcgggtgttggaagaagcaatgagataaagaacatcgcgtgacaaacga				
Db 123	TTTGAGAGATTTTGCTTTGGCCCAACCAACATGATTCAGAGATRCCATGTAAACAACCTGA				
Qy 3984	gtccgttgacacatcggctccatcgcgtcccgaaatctctcaaccggaanaatttcagag				
Db 183	GTACATGGACAATTTGGACATATGTCACAGATATCTTTCAACAGAGATTCCTCTGTAG				
Qy 4044	aaacccgacgttttcgatacgaatcatgctctctga				
Db 243	AAACCTGATGATTTGGCTATGAGATCATCTTTTGA				
RESULT 11					
LOCUS	B27171	464 bp	DNA	GSS	13-OCT-1997
DEFINITION	T2210TF TAWU Arabidopsis thaliana genomic clone T2210, genomic				
ACCESSION	B27171				
VERSION	B27171.1	GI:2513137			
KEYWORDS	GSS.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				

REFERENCE	TITLE	JOURNAL	COMMENT
1 (bases 1 to 464)	Rounsley, S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and Venter, J.C.	Use of a BAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing	Unpublished (1997)
Other_GSSs: T2E10TR	Contact: Steve Rounsley	Department of Eukaryotic Genomics	The Institute for Genomic Research
		9712 Medical Center Dr., Rockville, MD 20850, USA	Tel: 301 838 0200
			Fax: 301 838 0208
			Email: rounsley@tigr.org
			Seq Primer: M13-21
			Class: BAC ends
			High quality sequence stop: 464.
FEATURES	source	location/Qualifiers	
		1..464	
		/organism="Arabidopsis thaliana"	
		/strain="Columbia"	
		/db_xref="taxon:3702"	
		/clone="T2E10"	
		/clone_11b="TAMU"	
		/sex="hermaphrodite"	
		/note="Vector: BelongACII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"	
BASE COUNT	119 a	87 c	115 g 143 t
ORIGIN			
Query Match	3.9%;	Score 157.2;	DB 84; Length 464;
Best Local Similarity	64.5%;	Pred. No. 5.2e-23;	
Matches 320;	Conservative	0; Mismatches 108;	Indels 68; Gaps 3;
Oy 3002	aatttgatttcaggtggtgatagttaaccctggagcaatagctggtgagttgctgcaggt	3061	
Db 37	AATTAATTTAAAGTGGGTAGTACTCTACTGGAGCCATTGGGGAGGAGTTCCTGCTGT	96	
Oy 3062	gctccttgctcttcgctcctcctgcataagaccttctgtgttggtggcagaagaacaa	3121	
Db 97	GCTGCTTACTACTATTCTCTGCCCTCTTACTGTTGCTTGCTGTGGCTTAAAGAAACCT	156	
Oy 3122	ctagataattctctcgagtgctcggtgagttattatttcgcattagttctgtctctag	3181	
Db 157	CAAGAAATCTCTCTTGATGTCTCTGTAAAGTCACT	205	
Oy 3182	ccagcaatttggtttgctggagaaagtattggaacacactgttaatgaataatcaatacta	3241	
Db 206	CAACCTTGTTCATTTCAGAAATGGAA	235	
Oy 3242	gtcatgttttttaagtccaacacctttagtaaaatctcgattccaatactctatag	3301	
Db 236	GTAATTAGTACTTAATTTTTCAGTTCTGTG	268	
Oy 3302	cagccgaagaagatccagaagttcatctcgtggagacagctcaagaggtttcttcggggagc	3361	
Db 269	CAGCGGAAGAGGACCTGTAGGTTCACTTGGGGCACCTTAAACGGGTTCTCTACGGGAAAC	328	
Oy 3362	tacaagtgaggagtgatggtttagtaacaagaacaatttgggcagaggtggtgttggga	3421	
Db 329	TTCAAGTAGACACTGATAGCTTCAGCAACAAATAATTTTGGCCGAGGTGGGTTCGGA	388	
Oy 3422	aagttccaagggacgcttggcagacagcaactctgttctgttcaagaagactgaaggaag	3481	
Db 389	AAGCTCTCAAAATCCGCTGCTGGATGAGAACTGTGTGCAGTCAACAGCGCTTAAAGAAAG	448	
Oy 3482	agcgaaactccaggtgg 3497		
Db 449	AACGAAACCCAAAGTGG 464		

RESULT 12	
AA738545	
LOCUS	AA738545 1435 bp mRNA EST 17-AUG-1998
DEFINITION	SRRLX3 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
	SRRLX3 similar to protein kinase, highest similarity to
	receptor-like protein kinases, mRNA sequence.
ACCESSION	AA738545
VERSION	AA738545.1 GI:3421470
KEYWORDS	EST.
SOURCE	sorghum.
ORGANISM	Sorghum bicolor
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
	Poaceae; Sorghum.
	1 (bases 1 to 1435)
REFERENCE	Annen, F., Chang, J., -L., Paterson, A. H. and Stockhaus, J.
AUTHORS	Characterization of 14 different putative protein kinase cDNA
TITLE	clones of the C4 plant Sorghum bicolor
JOURNAL	Mol. Gen. Genet. 259 (1), 115-122 (1998)
MEMLINE	98409267
COMMENT	On Nov 29, 1993 this sequence version replaced gi:636085.

Insert Length: 1900 Std Error: 0.00.

FEATURES	SOURCE
Location/Qualifiers	1. 1435
/organism="Sorghum bicolor"	
/strain="cv. TX430"	
/db_xref="taxon:4558"	
/clone="SbRLK3"	
/clone_1lb="Sorghum bicolor cv. TX430 leaf"	
/dev_stage="green"	
/note="Organ: leaf; Vector: Lambda ZAP II; Site_1: Eco RI, Site_2: Xho I; leaf cDNA library from green leaves, unidirectionally cloned"	
BASE COUNT	365 a 300 c 389 g 381 t
ORIGIN	

	Query Match	Similarity	Score	157.2	DB	38	Length	1435
	Best Local	Similarity	70.5%	Pred.	M.5.3e-23			
	Matches	210	Conservative	0	Mismatches	86	Indels	0
								Gaps
QY	3784	ttgattggccaacgcggaagaagaatcgcgtcagctcagctcgaagttgtgtcttaacctac	3843					
Db	800	TGATTTGGCTCTCCAGGAAGCGAGTGCGCTATATAGGCACACCTCGTGGACGAGATTTTGC	859					
QY	3844	atgatacactcgcgcgtatccgaagatcattccaccggaagcgttaaaagcagaacaacccctcctag	3903					
Db	860	ACGACACCTCGCAATCCTAAGATTATACATCGTGACGTCGAAGCTGCGCAATGCTTGGCTTG	919					
QY	3904	acgaagaattcgaagcgggttgttggagaattcgggtttggcaagcctaatygactataaag	3963					
Db	920	ATGAAGGTTTTTAACCGGATGTTGGATTTGGGCTTGCCCAAGCGTGCGATGATACAA	979					
QY	3964	aacatcacttggcaacagcagatccgcgtggagacatcgtgtacatcgctccgaataatctct	4023					
Db	980	AACAATCTGTGGCTGCTCAGTGTCGCGGAACATATGGGTACATATGCCCTGAAATTTTGT	1039					
QY	4024	caacgggaaatctttagagagaacaccgacglttccggatcagggaaatcatgctcttaga	4081					
Db	1040	CCACTGCGGAAGTCATCCGAGAGAACCGAATGTTTGGTTACGGCAATAAAGCTTTCGA	1097					

RESULT	13
LOCUS	AI896277
DEFINITION	AI896277 555 bp mRNA EST 27-JUL-1999
ACCESSION	EST2851720 tomato callus, TAMU Lycopersicon esculentum cDNA clone
VERSION	LOC145, mRNA sequence.
KEYWORDS	AI896277
SOURCE	AI896277.1 GI:5602179
ORGANISM	EST.
REFERENCE	tomato.
AUTHORS	Lycopersicon esculentum
TITLE	Euphytica; Viridiplantae; Streptophyta; Embryophyta; Tracheophyt
JOURNAL	ephalycota; Spermatophyta; Magnoliophyta; eudicotyledons; co
COMMENT	eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanu
	Lycopersicon.
	1 (bases 1 to 555)
	Alcala, S., Vredalov, J., White, R., Matern, A.L., Vision, T.,
	Holt, I. E., Liang, F., Upton, J., Craven, M. B., Bowman, C. L., Ahn, S.,
	Ronning, C. M., Fraser, C. M., Martin, G. B., Tankeley, S. D. and
	Giovannoni, J.
	Generation of ESTs from tomato callus tissue
	Unpublished (1999)
	On May 18, 1998 this sequence version replaced gi:3137451.

Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES	SOURCE
location/Qualifiers	1. 555
/organism="Hyocrepisicon esculentum"	
/cultivar="TA496"	
/db_xref="taxon:4081"	
/clone="CLEC14U5"	
/clone_1lb="tomato callus, TAMU"	
/tissue_type="callus"	
/dev_stage="25-40 days old"	
/lab_host="XLI-Blue MRF"	
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2: XhoI; supplier: Giovannoni laboratory; clec - Copyedons of seedlings 7-10 days post-gemmation were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"	
BASE COUNT	154 a 90 c 148 g 163 t
ORIGIN	

Query Match	3.7%	Score 151	DB 62	Length 555
Best Local Similarity	71.3%	Pred. NO. 1e-21		
Matches 199	Conservative	0	Mismatches 80	Indels 0
OY	3784	ttgatgycgaacgcggaagaaatgcgtacggtcagctcgcaggttlylcttacctac	3843	
Db	275	TGGCTGTGTAAGCGCGAAAGAAATAGCAGTGTAGTACGAAGAGGGCTGTATTATTCG	334	
OY	3844	atgacacgtcgcgtccgaaagatcattaccgcgtgaagctaaagcgaacatccttag	3903	
Db	335	ATGAACAATGTGACCCCAAAATTATTCATCTGTGATGTAAAGCCCAACAATTTGTGG	394	
OY	3904	acgaagaattcgaagcgygttlytgaagatttcggttlygcaagctlaatyactataag	3963	
Db	395	ATGGGATTTTGAAGCAGTGTGTGGAGATTTTGGTTGTGCAAAACCTGTGATCACCGGG	454	
OY	3964	acaatccgtgaaacaaagcagctccgtggagacatggtgcacatgcgcgaatatctct	4023	
Db	455	ATTCTCACTGAAGCATGCTGCTGAGGGGACACAGTTGGTCACATTTGCTCCAAAATATCTTT	514	
OY	4024	caaccggaaaattctcagaagaaaccgacgttttcgat	4062	
Db	515	CAACGGGTCAATGCTGTGAAGAAGACTGATGTTTGGTT	553	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:12:15 ; Search time 12463.9 Seconds

(without alignments)
-86.322 Million cell updates/sec

Title: US-09-180-798-22

Perfect score: 1106
Sequence: 1 tcgaccacacgctcgcgcga.....aaaaaaaaaaaaaaaaaaaaa 1106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hgt1:*
33: gb_hgt2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hgt3:*
42: gb_hgt4:*
43: gb_hgt5:*
44: gb_hgt6:*

45: gb_hgt7:*
46: em_hgt1:*
47: em_hgt2:*
48: em_hgt3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_hgt8:*
53: gb_hgt9:*
54: gb_hgt10:*
55: gb_hgt11:*
56: gb_hgt12:*
57: gb_hgt13:*
58: gb_hgt14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1106	100.0	1106	5	A67817	A67817 Sequence 22
2	924	83.5	981	5	A67819	A67819 Sequence 24
3	912.4	82.5	1063	5	A67825	A67825 Sequence 30
4	714.2	64.6	789	5	A67821	A67821 Sequence 26
5	330	29.8	894	5	A67823	A67823 Sequence 28
6	191	17.3	2089	5	A67827	A67827 Sequence 32
7	151.2	13.7	936	8	SB062279	U62279 Sorghum b1c
8	96.6	8.7	4604	7	LEBRPGENE	X95269 L.esculentu
9	84.2	7.6	1755	8	DCU93048	U93048 Daucus caro
10	84.2	7.6	1815	5	A67797	A67797 Sequence 2
11	79.8	7.2	8153	50	AC021198	AC021198 Arabidops
12	77.8	7.0	8440	7	AB010075	AB010075 Arabidops
13	77.8	7.0	118086	7	ATF6H1	AL021684 Arabidops
14	76.6	6.9	5033	8	INU77888	U77888 Ipomoea n11
15	75.8	6.9	142418	7	AP000815	AP000815 Oryza sat
16	74	6.7	6695	5	A67796	A67796 Sequence 1
17	73.6	6.7	101284	8	ATAC009991	AC009991 Arabidops
18	72.4	6.5	4081	5	A67815	A67815 Sequence 20
19	72.4	6.5	98471	8	AC012654	AC012654 Arabidops
20	71.8	6.5	100887	50	ATAC011620	AC011620 Arabidops
21	70.8	6.4	3176	5	E12705	E12705 Arabidops
22	70.8	6.4	3176	8	ATU47029	U47029 Arabidops
23	70.4	6.4	112529	44	AC015446	AC015446 Arabidops
24	69.6	6.3	134402	8	ATAC011765	AC011765 Arabidops
25	69.4	6.3	116944	8	AC008017	AC008017 Arabidops
26	67.4	6.1	2786	7	AB029327	AB029327 Nicotiana
27	67.2	6.1	88401	8	F23M19	AC007454 Arabidops
28	66.8	6.0	69142	7	AB012239	AB012239 Arabidops
29	66.4	6.0	96475	7	ATP17M5	AL035678 Arabidops
30	66	6.0	111945	8	ATF1C12	AL022224 Arabidops
31	65.6	5.9	82033	7	AB011476	AB011476 Arabidops
32	65.6	5.9	86139	7	AB028616	AB028616 Arabidops
33	65.2	5.9	2278	8	SBRLX1	Y14600 Sorghum b1c
34	64.8	5.9	108355	50	AC005957	AC005957 Arabidops
35	64.4	5.8	61712	7	AB017061	AB017061 Arabidops
36	64.2	5.8	133840	8	AC003981	AC003981 Complete
37	64	5.8	3573	5	A57133	A57133 Sequence 4
38	64	5.8	3368	8	U42445	U42445 Lycopersico
39	64	5.8	6471	5	A57130	A57130 Sequence 1
40	64	5.8	6471	8	U42444	U42444 Lycopersico
41	63.4	5.7	87434	7	AB010698	AB010698 Arabidops
42	62.6	5.7	23143	8	F15P11	AF160760 Arabidops
43	62	5.6	65889	50	AC006436	AC006436 Arabidops
44	61.6	5.6	101154	8	ATF1P2	AL132935 Arabidops
45	61.2	5.5	3979	5	A67428	A67428 Sequence 1

ALIGNMENTS

```
RESULT 1
A67817
LOCUS A67817 1106 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 22 from Patent WO9743427.
ACCESSION A67817
VERSION A67817.1 GI:4756639
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 1106)
De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
AUTHORS PRODUCTION OF APOMICETIC SEED
TITLE Patent: WO 9743427-A 20-NOV-1997;
JOURNAL CIBA GEIGY AG (CH)
FEATURES
source
location/Qualifiers
1..1106
/organism="unidentified"
/db_xref="taxon:32644"
142..798
/note="unnamed protein product"
CDS
/protein_id="CAB42244.1"
/codon_start=1
/db_xref="GI:4756640"
/translation="MASRNYRWELFPAASLTTLTALIHVYANSEGDALYALRRSLTDP
DHVLOSMDPTLVNPTCFHYTCNDNRNVTVDLSNSGLAPELGKLEHLYELEY
KNNIGTIPSELGNLKNLISLDLNNNTLGIYPTSLGKLSLVFLRLDNRLTGPIPR
HSRSQAFKVVYDVSSNDLGGTIPNGPFAHPILOWENNPRLEGPPELLGLASYDINCT
"
BASE COUNT 331 a 258 c 206 g 311 t
ORIGIN
Query Match 100.0%; Score 1106; DB 5; Length 1106;
Best Local Similarity 100.0%; Pred. No. 5.8e-255;
Matches 1106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ttgaccacgcggtcgctcaactcaataaagggaacacgaatcccaatttgcct 60
Db 1 tgcaccacgcggtcgctcaactcaataaagggaacacgaatcccaatttgcct 60
QY 61 ttctcctcttcttgcaaaaatttccttcttactctcaaatcccttctgcctct 120
Db 61 ttctcctcttcttgcaaaaatttccttcttactctcaaatcccttctgcctct 120
QY 121 cttaaacctcgaagctcacaatgctcgcgaacatcctggtgagctcttcgacgt 180
Db 121 cttaaacctcgaagctcacaatgctcgcgaacatcctggtgagctcttcgacgt 180
QY 181 tggtaaaccttaaccttagcttgaattcaacctggtcgaagcaaatccggaagagatgt 240
Db 181 tggtaaaccttaaccttagcttgaattcaacctggtcgaagcaaatccggaagagatgt 240
QY 241 ccttaacgctcttcgcgcgaagttgacagatccagacacatgctcctcgaagctggatcca 300
Db 241 ccttaacgctcttcgcgcgaagttgacagatccagacacatgctcctcgaagctggatcca 300
QY 301 actcttgtaacctctgtacacctggtgtccatgctcactgttaacaaagaacacgcgtacct 360
Db 301 actcttgtaacctctgtacacctggtgtccatgctcactgttaacaaagaacacgcgtacct 360
QY 361 cgtgtgatttggaaattcaaacctctctggaacatcttgagcctgagcttggaaagctt 420
Db 361 cgtgtgatttggaaattcaaacctctctggaacatcttgagcctgagcttggaaagctt 420
QY 421 gaacattacagtatctagagctctaacaaaaacaacatccaaagaactataccttcgaa 480
Db 421 gaacattacagtatctagagctctaacaaaaacaacatccaaagaactataccttcgaa 480
QY 481 cttggaacactgagaactctacagctggtgactgttaaaaaacaacttctcaagatga 540
Db 481 cttggaacactgagaactctacagctggtgactgttaaaaaacaacttctcaagatga 540
```

```
QY 541 gttccaccttcttggaaaattgaagctctggtcttcttaacggttaatgacaaccca 600
Db 541 gttccaccttcttggaaaattgaagctctggtcttcttaacggttaatgacaaccca 600
QY 601 ttgaccggtcacaatcctcagaagcactcagcgaacatcccaagccttaagtgttgagctc 660
Db 601 ttgaccggtcacaatcctcagaagcactcagcgaacatcccaagccttaagtgttgagctc 660
QY 661 tcaagcaatgatctgtgtggagaacatcccaacaacgaagaccccttgcacatccctta 720
Db 661 tcaagcaatgatctgtgtggagaacatcccaacaacgaagaccccttgcacatccctta 720
QY 721 cagaactttgagaacacaccgagatttgagagagcaggaattactcgtgcttgcgaagctac 780
Db 721 cagaactttgagaacacaccgagatttgagagagcaggaattactcgtgcttgcgaagctac 780
QY 781 gaacactagacactgaaacacactgycacaaacctgaaatgaagaattggggggtagacc 840
Db 781 gaacactagacactgaaacacactgycacaaacctgaaatgaagaattggggggtagacc 840
QY 841 ttgtaagaacacttcaacacattatcaaatatcatcttatgtatgaatgaatata 900
Db 841 ttgtaagaacacttcaacacattatcaaatatcatcttatgtatgaatgaatata 900
QY 901 ttragaataaacaacaaatgaagaatcgtaataatcatcatcgtgctcaattgag 960
Db 901 ttragaataaacaacaaatgaagaatcgtaataatcatcatcgtgctcaattgag 960
QY 961 aacttcgagctgtatgtaaaattcctaattgcgatttcgctactcgttaagtctgct 1020
Db 961 aacttcgagctgtatgtaaaattcctaattgcgatttcgctactcgttaagtctgct 1020
QY 1021 tgtggattcttgagaagtaacatttgatgtatgataagttgttgccttgcct 1080
Db 1021 tgtggattcttgagaagtaacatttgatgtatgataagttgttgccttgcct 1080
QY 1081 gcaaaaaaaaaaaaaaaaaaaaaa 1106
Db 1081 gcaaaaaaaaaaaaaaaaaaaaaa 1106
RESULT 2
A67819
LOCUS A67819 981 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 24 from Patent WO9743427.
ACCESSION A67819
VERSION A67819.1 GI:4756641
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 981)
De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
AUTHORS PRODUCTION OF APOMICETIC SEED
TITLE Patent: WO 9743427-A 20-NOV-1997;
JOURNAL CIBA GEIGY AG (CH)
FEATURES
source
location/Qualifiers
1..981
/organism="unidentified"
/db_xref="taxon:32644"
104..760
/note="unnamed protein product"
CDS
/protein_id="CAB42246.1"
/db_xref="GI:4756642"
/translation="MASRNYRWELFPAASLTTLTALIHVYANSEGDALYALRRSLTDP
DHVLOSMDPTLVNPTCFHYTCNDNRNVTVDLSNSGLAPELGKLEHLYELEY
KNNIGTIPSELGNLKNLISLDLNNNTLGIYPTSLGKLSLVFLRLDNRLTGPIPR
ALTAIPSLKVVYDVSSNDLGGTIPNGPFAHPILOWENNPRLEGPPELLGLASYDINCT
"
BASE COUNT 286 a 236 c 180 g 279 t
ORIGIN
```


[illegible]

Db	570	TTAAACAGCTTCTCCGGCCCTATTCGGAAATCATTTGGGAAAGCTTTCAACCTAGATT	629
QY	580	ttacggcttaatagacacgcatgtagccgggtccaatccttagagcactcagcgcaatccaa	639
Db	630	CTCCGGCTTAAACAAACAGCTCTCAGTCTGGGTCATTCCTATGTCAGTCAGCAATATTACT	689
QY	640	gaccttaagctgttagcgtctcaagcagatgattgtgtggagcactccaaacagga	699
Db	690	ACCCTTCAAGTGTTAGATCTATCAAAATACAGACCTCTCGGTTCAAGTTCCTGACAAATGCG	749
QY	700	cccttgctcacatctccttacagacattgagaacaac	738
Db	750	TCCCTTCACCTCTTCACACACCATCAGTTTGTCTATAATAC	788
RESULT	7		
LOCUS	SB062279	936 bp	PLN
DEFINITION	Sorghum bicolor leucine-rich repeat-containing extracellular		09-DEC-1996
ACCESSION	U62279		
KEYWORDS	U62279.1 GI:1710123		
SOURCE			
ORGANISM	Sorghum.		
REFERENCE	Sorghum bicolor		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	euphyllipotes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;		
JOURNAL	Poaceae; Sorghum.		
REMARK	1 (bases 1 to 936)		
	Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.		
	Isolation of a cDNA encoding a novel leucine-rich repeat motif from		
	Sorghum bicolor inoculated with fungi		
	Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)		
	97124217		
	Erratum: [[published erratum appears in Mol Plant Microbe Interact		
	1997 Mar;10(2):302]]		
	2 (bases 1 to 936)		
REFERENCE			
AUTHORS	Hipskind,J.D.		
TITLE	Direct Submission		
JOURNAL	submitted (26-JUN-1996) John D. Hipskind, Botany and Plant		
	Pathology, Purdue University, West Lafayette, IN 47907, USA		
FEATURES			
SOURCE	Location/Qualifiers		
	1..936		
	/organism="Sorghum bicolor"		
	/db_xref="taxon:4558"		
	/clone="SLRR"		
	12..755		
	/function="protein-protein ligand binding"		
	/note="leucine-rich repeat-containing extracellular		
	glycoprotein; contains six N-glycosylation sites		
	[ln(S/1)]		
	/codon_start=1		
	/protein_id="AAC49559.1"		
	/db_xref="GI:1710124"		
	/translation="MAPQFAAGFLTGLATFASCTEGDILYKORLAWEDNNVYL		
	OSMSTGATPCPTWPCVPCNNNNPVYRIVRAGNSGTPPLATLQNTQYELGNGI		
	NGSPETIGANTNLISLDLMDNLGELPTPLGTSYTRIRLRQNNTGTPSPFGK		
	LTSLSEKLSQENSUGALPASIGNIKALQFRLNDMLTGVPSKSPFLSTFGNLTLE		
	NTDRNNLDGRTSSGLRTATAIIDALKTA"		
	36..77		
	/evidence="not experimental		
	BASE COUNT 247 a 232 c 207 g 250 t		
ORIGIN			
Query Match	13.7%;	Score 151.2;	DB 8;
Best Local Similarity	61.4%;	Pred. No. 3.1e-26;	
Matches 243;	Conservative 0;	Mismatches 153;	Indels 0;
Gaps 0;			
QY	223	aactcgaagagagatgctctacgctctcgccggagatttcacagatccagacatgctc	282
Db	81	AACACGAGAGGTCACCTACTGTACAAGCAAGATTGGCATGGAGAGACCCAAACAACGCG	140

QY	283	ctccagagctcgggataccaactctgtttaactcgttaccgtgtccatgtcaactgtcaac	342
	11		
Db	141	CTGCAGAGCTGGGAATTTCGAGCGTTGCCAATCCCTGCACCTGGTTCATGTCCACTGCAAC	200
QY	343	caagacacaccgcgtcacctcgtgtgatttggagaaatcaaacctctcgtgacatctgcg	402
	11		
Db	201	AACAAATAACTTGTCTACCCGCGTGATTTGGGCATATGACGACATCTCCGGCTCTGCTT	260
QY	403	ccgtgagcttggagagcttgaacatttaacggttctcaaggcctctcaaaaacaaatcaaa	462
	11		
Db	261	CCGATCTTCACAGAAATTCAGAACTCTCGATACATCGAGCTGTAATGGCAACGGCTTAAAC	320
QY	463	ggaactataccctcgaacttggaaatctgaaatctgaaatctcaacatctgagatctgtacaac	522
	11		
Db	321	GGTTCGATTCGAGAAACACTAGGCAACCTGACTATATCTCATACAGCTTGAGATCTGGAGC	380
QY	523	aacaattcacaggagtagtctccacccttctcgtggaataatggaatctcgtcttctt	582
	11		
Db	381	AACCTTTCTACGGCAAAATCCCACTACGCTTGCTGTGTGACAGCCTGCGATATCTG	440
QY	583	cgagtaatgacacacgattgacggctccatccca	618
	11		
Db	441	AGGTTGATCCAAACAACTGACAGGGGCTATACCA	476
RESULT	8		
LEIRPGENE		LEIRPGENE	4604 bp
LOCUS		L. esculentum	LRP gene.
ACCESSION		X95269	
VERSION		X95269.1	GI:1619299
KEYWORDS		LRP gene; LRP protein.	
SOURCE		tomato.	
ORGANISM		Lycopersicon esculentum	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;	
		asteridae; Solanales; Solanales; Solanales; Solanales; Solanales; Solanales;	
		Lycopersicon.	
REFERENCE		1 (bases 1 to 4604)	
AUTHORS		Vera, P.	
TITLE		Direct Submission	
JOURNAL		Submitted (18-JAN-1996) P. Vera, Universidad Politecnica de	
		Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camilo	
		de Vera 14, E-46022 Valencia, SPAIN	
		2 (bases 1 to 4604)	
AUTHORS		Tornero, P., Mayda, E., Gomez, M. D., Canas, L., Conejero, V. and Vera, P.	
TITLE		Characterization of LRP, a leucine-rich repeat (LRP) protein from	
		tomato plants that is processed during pathogenesis	
JOURNAL		Plant J 10 (2), 315-330 (1996)	
MEDLINE		96367673	
FEATURES		Location/Qualifiers	
source		1..4604	
		/organism="Lycopersicon esculentum"	
		/cultivar="VFN8"	
		/db_xref="taxon:4081"	
		/tissue_type="leaf"	
		<1..345	
exon		/number=1	
gene		113..4355	
		/gene="LRP"	
		join(113..445,2381..2452,2565..2708,2928..2999,3081..3152,	
		4283..4355)	
		/gene="LRP"	
		/codon_start=1	
		/product="LRP protein"	
		/protein_id="CA64565.1"	
		/db_xref="GI:1619300"	
		/db_xref="SPTREMBL:O96447"	
		/translation="MEAVKQSLFKLGMGLAVAVAVKGSBDALVYLRSL	
		SDPGVLIQSDMDPLVNPCHWPHYTCNGQVTRLDGNSKISGILVEIKLEHLYL	
		ELKNNIGITPREJANLSLSLDLYNNNSGITPTPSLGAKILVLRILNDKNTGP	
		IPETKISISLKAIVDSSNDLCGTTPTSGPEEHIPUNNEHNPLDEPPELLGATADYT	

	BASE COUNT	ORIGIN
intron	346. .2380 /gene="LRP" /number=1	NCS"
exon	2381. .2452 /gene="LRP" /number=2	
intron	2453. .2564 /gene="LRP" /number=2	
exon	2565. .2708 /gene="LRP" /number=3	
intron	2709. .2927 /gene="LRP" /number=3	
exon	2928. .2999 /gene="LRP" /number=4	
intron	3000. .3080 /gene="LRP" /number=4	
exon	3081. .3152 /gene="LRP" /number=5	
intron	3153. .4282 /gene="LRP" /number=5	
exon	4283. >4604 /number=6	
BASE COUNT	1311 a 758 c 888 g 1647 t	
ORIGIN		

Query Match	8.7%	Score 96.6	DB 7	Length 4604
Best Local Similarity	69.1%	Pred. No. 3,7e-13		
Matches 132	Conservative 0	Mismatches 59	Indels 0	Gaps 0
QY 178	gcttcgttaacccttaaccttagcttggatlttaccctgctgcaagaacaactccgaagagat	237		
Db 158	GGTTTGTGGCAATTCTTCTTGCTTAACCTTGCGCTGTAAAGGGGAATTTCAGAAAGGGGAT	217		
QY 238	gctctcaagcctcttcgcgcgagatttcagacagatccaagacatgctccctccgaagctggat	297		
Db 218	GCATTGAGGCCCTTGCGCGAGCTTATCTGATACCGGGTAACGGTTACGAGACTGGAT	277		
QY 298	ccaactcttgtaatcctctgtacctcgtgttcacatgtaacacgaacaacgcgcgc	357		
Db 278	CCAAATCTTTTAAACCTTGCTGCTGCTTATGTCATGCTACCAAGGATATCAAGTT	337		
QY 358	actcgtatgga 368			
Db 338	ACTGCTGTGA 348			
RESULT 9				
LOCUS DCU93048				
DEFINITION Daucus carota somatic embryogenesis receptor-like kinase mRNA, complete cds.				
ACCESSION U93048				
VERSION U93048.1				
KEYWORDS GI:2224910				
SOURCE carrot.				
ORGANISM Daucus carota				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiales; Daucus.				
REFERENCE 1 (bases 1 to 1755)				
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.				
TITLE A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos				
JOURNAL Development 124 (10), 2049-2062 (1997)				

BASE COUNT	ORIGIN	FEATURES	source	TITLE	AUTHORS	REFERENCE	MEDLINE
506 a	347 c	407 g	495 t	LAERDEMOKVEVIAHODVELAPRTSGMILSTDINHAELESGP"	Submitted (12-Mar-1997)	Direct Submission	97313247
				1..1755	Location/Qualifiers		2 (bases 1 to 1755)
				"/organism="Daucus carota"			97313247
				"/db_xref="taxon:4039"			2 (bases 1 to 1755)
				34..1695			Schmidt,E.D.L., Guzzo,F., Toonen,M.A.T. and de Vries,S.C.
				/note="SERK"			Direct Submission
				/codon_start=1			Submitted (12-Mar-1997)
				/product="somatic embryogenesis receptor -like kinase"			Molecular Biology, Agricultural University
				/protein_id="BAB51708.1"			of Wageningen, Drentheaan 3, Wageningen 6703 Ha, The Netherlands
				/db_xref="GI:2224911"			
				/translation="MNRNSINILNWMQFDAYIDKRGVIMLTLEYLSNNISGSPISDSDG			
				NITNIVSYDILYMNFSFGPIPDPTGKTLRLRLRLNNNSGSPISGIPITTPSTVD			
				SNRRISGVPDNGSFSFLFTTYSFRNNINLDGFTGRCPSFPSPPTTPTSTVD			
				PEQNSPTGIALAGVAAAGALLFAPAPMAAPRRRRPRHFPDVAEEDPEVHIGDLK			
				FEPSLELVADTDTSTILGGGEGKVKYKGLADGSLVAVKRLREBRTGDEGLAQOTE			
				ALMSVAHRYNRLRLRGCPMTPTERLLVTPYMGNSVASCILRRQSPSEPPDMPKRII			
				ALGSRGSLYHLDHCDPKIITHRDYKANILILDEEFAVYGDGLARLDYDKDTFTVTA			
				VSGTIGYTAPEYILSTGSKSEKTDVEGGIMLELILGQAPFLATLADDDVIMLDWM			
				SLKLEKTEMLVDPDLQENNYIDIEVEQLIOLALITCQGSPPMRKSEVYRMLESGG			
				LAERDEMOKVEVIAHODVELAPRTSGMILSTDINHAELESGP"			

Query March	7.6%	Score 84.2	DB 8	Length 1755	
Best Local Similarity	52.1%	Pred. No. 3.5e-10			
Matches 188	Conservative	0	Mismatches 173	Indels 0	Gaps 0
QY 399	tgcgcctgagcttgggaagcttgaacatttacgatactagactctacaacaacaacat	458			
DB 78	TGATGCTACTACCTTGACAAATATATGGGGGCTCTTATGACATTTGAGCGTTTACGACATATACAT	137			
QY 459	ccaaggaacatactacctccgaactctgaagatctgaagatctcaagcttgatctgta	518			
DB 138	AAGTGGACCAATTCCTAGTAGTCTTGGGAACTGTGACAAATTTGGTGTGACTGGACCTATA	197			
QY 519	caaaanaaatcttaacgggatagtcccaactctcttgggaaaatgtagctctgctct	578			
DB 198	CATGAATAGCTTCTCTGGACCTTATACCGGACACATTAAGAAAGGTTTCAAGGCTAAGATT	257			
QY 579	tttaagcgttaatgaacacggaatlaccggttccaaatccctagaagcaccacggaatccca	638			
DB 258	CTTCGCTCTCAACAAACAAACAGCCTCTCTGGTCCCAATTCACATGTCACTGACATATATAC	317			
QY 639	agccttaaaagtgtgtagcgtctcaagaatgattgtgtgggacaatcccaacaagg	698			
DB 318	AACCTTCAAGTCCCTGATTTATCAACACATAGGCTATACAGGACCACTACCGGATATAG	377			
QY 699	acccttgcatacttcccttacaagaactttgagacacacccggaattggagggaccgga	758			
DB 378	CTCATTTTCTTTGTTTACACCTATCAGTTTGCCCAATTAATTGAATTTATGTGTGACCCG	437			
QY 759	a 759				
DB 438	A 438				
RESULT 10					
LOCUS A67797	1815 bp	DNA			
DEFINITION Sequence 2 from Patent WO9743427.					
ACCESSION A67797		PAT	05-MAY-1999		
VERSION A67797.1	GI:4756623				
KEYWORDS					
SOURCE	carrot.				
ORGANISM	Daucus carota				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				


```

RESULT_12
AB010075          ABO10075      84440 bp       DNA           PLN             20-NOV-1999
LOCUS              Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MPA24, complete sequence.
DEFINITION
ACCESSION          AB010075
VERSION            ABO10075.1   GI:2760171
KEYWORDS
SOURCE             HMG
                  Arabidopsi s thaliana (strain:Columbia) DNA, clone_id:Mitsui PI
ORGANISM           Arabidopsi s thaliana
                  Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumphylophyes; Spermatophyta; Magnoliophyta; euclcotyledons; core
REFERENCE          Arabidopsi s.
                  1 (sites)
AUTHORS            Nakamura,Y.
TITLE              Structural Analysis of Arabidopsi s thaliana Chromosome 5. IV
JOURNAL            Unpublished (1998)
REFERENCE          2 (bases 1 to 84440)
AUTHORS            Nakamura,Y.
TITLE              Direct Submission
JOURNAL            Submitted (07-JAN-1998) to the DDBJ/EMBL/GenBank databases.
                   Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2; 1532-3, Yena, Kisarazu, Chiba 292, Japan
                   (E-mail:y-nakam@kazusa.or.jp, Tel:+81-438-52-3935, Fax:+81-438-52-3934)
FEATURES
source             Location/Qualifiers
                   1..84440
                   /organism="Arabidopsi s thaliana"
                   /strain="Columbia"
                   /db_xref="taxon:3702"
                   /chromosome="5"
                   /clone="MPA24"
BASE COUNT         27069 a 14857 c 15226 g 27288 t
ORIGIN
Query Match        7.0%; Score 77.8; DB 7; Length 84440;
Best Local Similarity 52.3%; Pred. No. 1.2e-08;
Matches 172; Conservative 0; Mismatches 157; Indels 0; Gaps 0
OY               384    ccctcctcggaactctgcgacctggagacgttgaaaccttaaccagtatcatgagct 443
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db                35029 CTTCTCCGGTCATTAACTTGGAAGCTAGGAAAGCTATCGAGTTTAAAATCAATGGAATT 35088
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY               444    ctcaaaaacaacataccaaggaaactatacccttcggaacttggaaactgaaatctgaagaattcat 503
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db                35089 GTCTAACACAATGTMTACCACCGAGAGATTCACGAGGAGTTGCCAGAATTAAGAGATCTCAC 35148
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY               504    cagcttgatctgtcacacaacaactcttacaggatagttgccactttcttggaaaaatt 563
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db                35149 GCTTTGGAATCTTCCTCCGTAAACAACACTTCACGCAGAGATACCGAGATTTCATCGAGATT 35208
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY               564    gaatctctgctctttcttacgacgttaaagcaaacgcatigacacggtgcacatctctaagaca 623
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db                35209 GCCCGAGCTTTAAATGTGTACAGCTTTGGGAAACAATTTCACCGGAAGCATCCCCAGAGA 35268
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY               624    ctcaagcgaatccccaaagcctttaagtttgttagcgtctcaagaatatgtattgtgtggac 683
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db                35269 ATTATGAGATAAAAAAGGTAAACTTAATCTCGTCGATCTCTTCCAATAAGCTGACCGGAAC 35328
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY               684    aaaccaacaagaagacccttctgtcata 712
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db                35329 TTITACCGCCGAACATGTCTCCGGTAGACA 35357
                    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT_13
ATF6H11/C
LOCUS              ATF6H11     118086 bp       DNA           PLN             22-APR-1998
DEFINITION         Arabidopsi s thaliana DNA chromosome 5, BNC clone F6H11 (ESSAITI
```

accession	AL021684.1	GI:2827698	project . AL021684
version	AL021684.1		
keywods			
source			thale cress.
organism			Arabidopsis thaliana
reference			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae; Capparales; Brassicaceae; Arabidopsis.
authors			1 (bases 1 to 118086) Bevan,M., Brandt,P., Dose,S., Jarke,D., Scharfe,M., Schon,O., Hoheisel,J., Mewes,H.W., Mayer,K. and Schuelter,C.
journal			Unpublished
reference			2 (bases 1 to 118086) EU Arabidopsis sequencing project.
authors			Direct Submission
journal			Submitted (20-Apr-1998) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk
features			location/Qualifiers
source			1..118086
			/organism="Arabidopsis thaliana"
			/variety="Columbia"
			/db_xref="taxon:3702"
			/chromosome="5"
gene			1205..5836
			/gene="F6H11.10"
			join(1205)..1314,1384..1598,1651..1669,1746..3079, 4085..4615,4571..4960,5255..5836)
CDS			/note="similarity to predicted protein (chromosome IV), 229203.1238035, 226379, 226007"
			/codon_start=1
			/product="predicted protein"
			/protein_id="CA16672.1"
			/db_xref="GI:2827699"
			/translat="MGTYGSSOGNGYLTGSELAHVAGHGRDELITGYSLKSKPKKE ITGPQKRSLFLICFALITISPSKPTQPILEPFLPHKTVYVFLNFSQIHILG SVPRQSGBALIMKKDPVFRMPFQPSKNGGFDHIDLHSAVDGVGGDGGFTLE VWTDKRLDKEMISELEPELPSVFICISLEPQDVTVLTQTERENILKWN IGCHCPTTMOELMDWDLVTPNKTLOHLYTWEQKVVLMKKRSEVOGRAIEILGTLR KKSGAKYHALSELQVYMAHAALAKTVVDGSEVYISLLEPFLSHVGEALAIYV NLELDSQKAGLMOFAPVNSLWMLNDGSIETETINCARLIGLVEKEGFRALVSSH LIVGLMRLVYRDRRRNGVSPALITLKSYSVHKQVRLRYRIGAEVQALVDVYPCJDI LESAFLVDSICLSESERIALKDSYNTPIHTVYLLAKYSEKTAQNTAISVSKLAS ECCSLAYEVGLAKLLVLTQSGDPLMKRQSAELKLCISLHYSQMTISYMMALNA STANSUKLDEEDDDHKLNQPSSEVSQDSQRYGKEDIISLITLTLPQVGAASF SOPSSPDLQERNQTPHPREDEDLAGIKLNFQFAELIGRFFRGISKLGNLPSVET IANSNLQSLSEDDVPKQVDVYGVTEELVAFVKDLAHPEITWDFPLPDQDSFVKE PFVSLTAAPVIGLVGSVILNCLFNLOSDEFLDAQVEMHALAVERLAPLASLALRIE LCPEMTENGSFWRITVYFLVHLNPKLSKDHALLSTPOVLEKSMQLSOGLCRKSLV SEANATVYVEPLTVTPPSPETRAKYIYNPVSSSEYEDKRPITSEKIQLYDYSVIEKS TSAASSRFINVOYDDDDADMDLNDETSSVSAIGRSTNRPFEDEEDVSFSD LEEDDNGDVKVSYNKLTINSQSDSDKSPDWVQLKEVKEKKKNMDLVDAV"
exon			1205..1314
			/gene="F6H11.10"
			/number=1
intron			1315..1383
			/gene="F6H11.10"
			/number=1
exon			1384..1598
			/gene="F6H11.10"
			/number=2
intron			1599..1650
			/gene="F6H11.10"
			/number=2
exon			1651..1669
			/gene="F6H11.10"
			/number=3

QY 624 ctcacggaatcccaagctttaaagtgtgacgtctcaagcaatgattgtgtggac 683
 DB 84538 ATTAGGAGAAACGGTAAACTAATCTCGATCTCTTCAATAGCTGACCGGAC 84479
 QY 684 aatcccaacaagcagccttgctgcaca 712
 DB 84478 TTACCGCGGACATGTGCTCGGTAACA 84450

RESULT 14
 INU77888
 LOCUS INU77888
 DEFINITION Ipomoea nil receptor-like protein kinase (lnrpk1) gene, complete cds.
 ACCESSION U77888
 VERSION U77888.1 GI:1684912
 KEYWORDS Japanese morning glory.
 SOURCE
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
 REFERENCE 1 (bases 1 to 5033)
 AUTHORS Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajeevan,M.S.
 TITLE Identification and Preliminary Characterization of an Unusual Leucine-rich Repeat Receptor-like Protein Kinase from Morning Glory (Ipomoea nil)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 5033)
 AUTHORS Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajeevan,M.S.
 TITLE Direct Submission
 JOURNAL Submitted (12-NOV-1996) USDA, ARS, Appalachian Fruit Research Station, 45 Wiltshire Road, Kearneysville, WV 25430, USA

FEATURES
 SOURCE
 1..5033
 /organism="Ipomoea nil"
 /cultivar="Violet"
 /db_xref="taxon:35883"
 /tissue_type="young leaf"
 237..251
 681..695
 1144..1456
 /gene="lnrpk1"
 /join(1144..4108,4202..4566)
 /gene="lnrpk1"
 /note="leucine-rich repeat; protein kinase homolog"
 /codon_start=1
 /product="receptor-like protein kinase"
 /protein_id="AAB36558.1"
 /db_xref="GI:1684913"
 /translation="MKAVNTFLSLCSSTSYTAAPALNSDGAALLSLFRHWTPIPSD ITQSNWADSTPCSMGVECDRQFVTLNLSYGISSGEGPISHLKIKKVLISGN GFPGSIPQLNCSLLEHIDLSNFSFTGAPDITGALQNRNLSEFNSIGTFPESL LSPLELFTYFTGNGLNGSIPNSGNSSELTTLDNNGSPGSPSNGITGLDLY LNDNNVGLTFLNTLNLENTLYLDVRNNSLGAIPDLFVSCQIDITISLNNQFTGL PPGNGCTSLRFGAFSCALSGPIPCFQGLTQITDITLYAGNCRHSRIPPELCKGSM IDIQDQNLGSGTPELGHMSQQLYHLYTNNISGEVPSYTIKISDSISDLYQNRU SGELPDVNTLELQVLSALYENHTGVIPODLGANSLEYLDLPFEKONLTFPNCIS OKTKRLRLGYNVLSGVSDLGCGSTLERLLENNLRGLDPFEKONLTFPNCIS NNTFGPIPPSGLNKVAYLSSNOLSGSPPELISGLVLEHNLNSHLNKLIPSE LSNCHRLSELDSASHNLNGSIPSTLGSITTELTKLSGENSGGIPSTLSQSKLNL OLGNLTLADGIPVGAQLQALRNLNSKLNKGQPIDLGKLMLELDVSHNLSGTL RYLSTIOSLFTINSHNLSFGPVPSPKFLNKGSPSTFSGNSDCLNCPAGLACPES SLIRPCNNQMSNGGSLTGIMIVIGALLFTICLPISAPLFIKCKSVOCISATSA OEGDSILNKYLEATENLNDKYVIGKAGHGYIKATISPRKYAVAKKLYKQVNSA SWARLETIGVYRHNLLKLEFWLKRKIGLITTYMENSPLVILHEINLPKPLWS TRHNIAVTAAGLAVLHFDQPAIVHDIKPMNLILDSLEPHISDGLAKLIDQSAT SLPNTVOGTIGYAVAPENAFITVRSRDYSGVLELITRKALDPFENEDTIV GWNRVMTQGTGIQIKIYPSLDELIDSVMEQVETALSLALCAKEVDRKRPMTADV VKQLTRMSIRSRSSSVRNKSK"

BASE COUNT 1365 a 1060 c 1017 g 1591 t

ORIGIN

Query Match 6.9%; Score 76.6; DB 8; Length 5033;
 Best Local Similarity 49.1%; Pred. No. 2.3e-08;
 Matches 202; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 285 ccaagagctggatccaactctgttaactcctgttaccctgttccatcactgttaacca 344
 DB 1281 CCAGAGCTGGAAATGCTTCAATTCACATCTCTTGTATGCTGGAGTGAATTTGACAG 1340
 QY 345 agacaacgcgctcaactcgttggattggaaattcaaaccttcggacatctggcc 404
 DB 1341 GAGACAAATTTTATATCTCTGAACTCTCTCTTAATGAAATCTAGGCAATTTGGGCC 1400
 QY 405 tgaacttggaaagcttgaacattacagatctagaactcaaaaacaacatccag 464
 DB 1401 CGAAATCTCGCATTTGAAAGCATTTGAAAGAGTGTCTCAGTGCAGCAATGGTTCTTGG 1460
 QY 465 aactatccttcggaacttggaaatcgaagaatcgaatcgaatcgaatcgaatcga 524
 DB 1461 CTCAAATTCCTTCCAGCTAGGCAATTCAGTCTTCTTGAACACATGATCTGCTCCA 1520
 QY 525 caatctcaaggaatagttcccaattcttctgggaaattgaagtcctgtctttttag 584
 DB 1521 CAGCTTACTGTAATATCCCTGACACCCCTTGAGCTTTGCAGAAATTTAAGAACTTAG 1580
 QY 585 gcttatatcaaacgattgacacggtccaatccatcagagcaactcaggaatcccaagc 644
 DB 1581 CCGTCTCTTAAATTCCTGATGGCCATTTCCGATCTTCAATTCACATTT 1640
 QY 645 taaagtgttgaactcgaagcaatgattgttggagcaatcccaaca 695
 DB 1641 AGAACTGTATTATTCACGTGCAATGCTTAAATGTTCAATCCCTTCAA 1691

RESULT 15
 AP000815/c
 LOCUS AP000815
 DEFINITION Oryza sativa genomic DNA, chromosome 1, clone: P0003H10.
 ACCESSION AP000815
 VERSION AP000815.1 GI:6498418
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa (cultivar: Nipponbare) DNA, clone: P0003H10.
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; Oryza.
 1 (bases 1 to 142418)
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC clone: P0003H10
 Published only in Database (1999) In press
 2 (bases 1 to 142418)
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Direct Submission
 Submitted (30-NOV-1999) to the DDBJ/EMBL/GenBank databases, Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail: tsasaki@labr.affrc.go.jp)
 URL: http://www.dna.affrc.go.jp:82/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT
 The orientation of the sequence is from T7 to SP6 of the PAC clone. Genes were predicted from the integrated results of the following: GENSCAN 1.0, BLASTN 2.0, BLASTX 2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against the non-redundant database NRP (PIR, SWISSPROT, GENPEPT, PDB) from MAF DNA bank and the cDNA sequence database at RGP. Protein similarities of the coding regions were searched against NRP with BLASTP 2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone id.

Detailed information on assemble quality together with annotation of this entry at <http://www.dna.affrc.go.jp:82/genomicdata/genome1>

FEATURES	source	nished.html	Location/Qualifiers
CDS			1..142418 /organism="Oryza sativa" /cultivar="Nipponbare" /db_xref="taxon:4530" /chromosome="1" /clone="P003H10" join(2901..2951,3129..3317,3521..3850,4405..4638,5069..5509) /note="Similar to Arabidopsis thaliana DNA chromosome 4, BAC clone F22K18 (A035356)" /codon_start=1 /protein_id="BAA87822.1" /db_xref="GI:6498419" /translation="MGEPNQRRLRYQWOGSNKPLCGGRILFGPDAGSLSTVLIVAPLVGACCCITKANSISEKQVGLPYLIATIVGLADLAFVLTMTSSRDGIPRNARPPEGGGDEGVAGDYTPPSAEWTAAPHLRLPRKDYVNGCVYKVCDCILLYRPPRASHSICNNKVRKRDHRCPPWGCIGLRNRPFLFISTILCYVVPYSMTIVARHGDGSLKSMAGEPLSVLYITVFVSVFVGLVFLHLYLSTQITENFRIYDKKENPVRGALISNAIEVFCAGIPSNMNPWAPPLEPDPVSGQLPRNGADLTGGVKEKVDLEWRNGGILPAILRGIDYDEMKNDSVHIKDRGAAPAPDFVAGRWHNDECESTPAAVSHVNSERL" join(9165..9635,9735..9799,9992..10319,10422..10520,10743..11006) /note="ESTs D47452(S12946), C23573(S10086), C19173(E10057), A0081294(E10057), A0032453(S10086) correspond to a region of the predicted gene.; Similar to PN134. (037437)" /codon_start=1 /protein_id="BAA87823.1" /db_xref="GI:6498420" /translation="MASSAMEILSNPAAVRGLSAKPRVSSRRIVRFVSSAAPPAARCPKRRKGTIEOETLLTPRYTDDDEMRFPNAENKQLNOEEDLLOEKTIDVQITFVRNPERKADAKMEGLROIPEVLEERCTAFESGLIKELGRRLKTNPVVAEIFSLMSRDEARHAGFLNGLSDFNALDGLFATARKTFEKFPIIYATYLSKIGYWRVITIFRLKANPQVAYPIKFEYENCCODENRGGFSAALAAQDFLMDKAKLMSRFCLSYVYMYLNDQCTREYEGIGLDEKFEHDVLIETNRNTRAFEPVLDVNEPERKLDRAVEINRKIIAIGESDDIPLVKNKRIPIHVAALVSEIIAATIMPEI ESGVDAEPPEPOLVY" join(13373..13457,15836..15914,17337..17477,17560..17654,17726..17860,18400..18587,18681..18749,18826..18916,19028..19162,19323..19379,19349..19441,19450..19509,19523..19585,19589..19701,19800..19941,20057..20294,21339..21586) /note="ESTs A0081256(C53656), A0081257(C53656), A0032236(R3784), A0081351(E61905), A0081361(R3784) correspond to a region of the predicted gene.; Similar to alien-like protein. (AC005623)" /codon_start=1 /protein_id="BAA87824.1" /db_xref="GI:6498421" /translation="MEDYGFESDDEPEDVDIENQYNSKGVETDPGALAGFDVVRMEPRAEKGFALQOTVLYKLGKREKMDAYEMLTYSKSAVTRYSEKCNINDVYSSASQNSFSLQEPYOTIKALEAKNERLAKTILKCKIMPDMGEGRSKILKEHNSCOREDSDDOKGTQLEIYATLEIQITETKNNKLKLEYITKALSKAI PPHRMGILRECGGKMAERQWADATDFEAFKNDENGNPRRIQCLKLYLALANLMESEVNPFGDEAPRYKNDPEILAMTNLIAQYONDIMEEKLKSNRRIMDDPIR NYIEDLKNIPTOVLKILKYTRIRIPIFSOKKCMKISNMEIWTILVLFHTIYTVRAIGSCFKYSIKLDFLSCQRTVOLKSCALNADPMLFTEVNLIKSYVOALASI VALSMCELNPERKDEVDOLVSLIDNRIGOHIDOVKLELRDRETPSHROVSAE EHLNVOQSVNRRIACGCTAGTCLHLYLIMMGAFCEPLNLSANSNSETILASM ARNFWFCNHCSTKFWYGREVWPRVRIAPGTRTIANIVLPSSDSVANDSLSNQNG KESPLSTYTDQALNLOREHTTILSDVINLSSESYRLAPSSHLGHPVYITIVDNN I" join(22475..22550,23332..23705) /note="ESTs A0030013(E50493), A0081341(E50493) correspond to a region of the predicted gene.; Similar to O.sativa gene encoding calmodulin. (Z12828)" /codon_start=1 /protein_id="BAA87825.1" /db_xref="GI:6498422" /translation="MADQLTDQIAEFKEAFSLPDKDGCITTKELGTVMRSIGNP TEALQDMINEVDADNGTIDIPPEFLNMARKMKMDTVSEELKAPFVFKDQNGFIS
CDS			AAELRHYMTNLGKELDEEVENIREADVDGQGINYDEFKVMAR" complement(join(24778..24897,25615..25722,25808..25993,26090..26177,26435..26638,27519..27581,27708..27778,28329..28389,28512..28552,29058..29225,30476..31136)) /note="hypothetical protein" /codon_start=1 /protein_id="BAA87826.1" /db_xref="GI:6498423" /translation="MMFQTSSTDARTHSGSDRGSAGRIACRGTTRSAAPGSGREAT TRGHALPEPPASPRLYVGRGSRSPSPPLPYTARCEEHRHRRCSLPPR RORRGAPPPDPASPCPDLATLGVGRMKMAAEVGVAAEEEMARAROLRG GGVYVVAHGRKAREERRCGGEERSGDEGRVREEDKAEPPYGRNKGEGTGE TDERMNGHSHKRSITLMQITKPPARRGAVLAHCGDDRRRCRCACVARRROE ALPIQSLPILFIMRDLKAYAKEQDIGYAGVASCITQYASVCKREHVALISL VTSRRAIALVVGPAIGFUSQETLHMHODKEVIDLXADNTSDIGETKSGSRM GHTSRLKNALVQMSAITLVCESLHDTAYLEIFSLMAVSSRKYRGISFTSDVGVILA ISGFGVLYOLAIPLIAKYGP IKFEPFAVYLSILLSTYFPMNLVLELKVLINI ASLKNEARKKEMTOSISLDAGFGRKSTLGCSTYVKGQSGSKP" complement(join(34468..34544,37202..37385)) /note="hypothetical protein" /codon_start=1 /protein_id="BAA87827.1" /db_xref="GI:6498424" /translation="MAAGDKAGDDDAAAAAPLTVSAAGRRRCPCLTBERCRADA GIPLNFEYIMVYICSEFPRRYRPAKKICDEFSKAMAHGHR" join(40347..40358,40468..40680) /note="hypothetical protein" /codon_start=1 /protein_id="BAA87828.1" /db_xref="GI:6498425" /translation="MHPKHLVAAATVILVQSPLEPSSFSNNRRRRROCRPTIAT ADGSTRVLPDPDGGKGFRRMRKEGG" complement(join(42025..42083,43481..43661)) /note="hypothetical protein" /codon_start=1 /protein_id="BAA87829.1" /db_xref="GI:6498426" /translation="MAAGDKAGDDDAAAAAPLTVPAAGRRRCPCGTRKERCENAHG IPYLNFEYIMVYICSAVNRVENTEPTSTSVS" join(47081..47084,47194..47313,47469..47584) /note="hypothetical protein" /codon_start=1 /protein_id="BAA87830.1" /db_xref="GI:6498427" /translation="MSRGRARATACGSYRCYAAEAEQEDSARVALLLRDSP AATPPLIARRIPHDGPPLEREDLSGPTLD" join(48412..48822,49072..49179,49283..49438,49876..50091,50202..50411,50488..50625,50842..50943,51105..51170,51274..51358,51454..51531) /note="Similar to Arabidopsis thaliana chromosome II BAC T27A16 sequence; hypothetical protein. (AC005496)" /codon_start=1 /protein_id="BAA87831.1" /db_xref="GI:6498428" /translation="MVLPLPVSRCRPGSPAPRSRRFLDPGGGNGVDGAGVYF RRLRLGTVDSNTSSSSSRGRGRDRAHGGYGDGDADALASVRLRLSSAODD AAEEBAEDDEGQPKRMAYVLCFSAFLCNDNRVMSIALIPMAKSGWNPQVTL IOSFEFMYDITQJAGCIWADTVGKLVLFQVYINRMAIALPFAKGLPLVTVR AFMGVGEVAPANNILSLMNPVSESRSLAYVSGMYLSTYGLAFSPILHNHGM PSVYFSGSLGGEVFWFSTWASKAVSSPLEDGIASKEEKLITSGTGGEPKLPACLI LSKRPVVALYISRCCHMWGTFILLTMWRYYNQVYKIRLESGLFCVILPMTAVASN FEGMADTVLSRGISVTVYKRLMSITFLPAPFLIOLSHIDSPMAAVLCMCSQSD AFOSGSLVNHQDIGPRAGVGLSLSTAGVLAGVETAAVTGILLDGHSDVDVFKSV VLYLVGLVWNLFFSTGEKIID" complement(join(52616..52685,53228..53299,53395..53466,53657..53800,54663..54734,56026..56234)) /note="ESTs C97644(C60871), A008165(C60871) correspond to a region of the predicted gene.; Similar to L.esculentum LRP gene. (X95269)" /codon_start=1 /protein_id="BAA87832.1" /db_xref="GI:6498429" /translation="MAAWSPALAAVLLAAVASANSSEGDALYALRALADPRGVILQ

CDS

SRDPTLVNPTWETWYTCDBRAGVTRLDGNSNLSGHLAPELGHLEHLOYLEYKNNTIO
GTPAPELGSLKNIISLDLYNNNTGTIPRELGLSLIYFRLNDSLSNGPIPRDLAKI
SSLKYIDVSNNDLCGTTPTSGPFEHPIPLNFPDKNPRLDGPFLQGLATIDYNC"
59608..59826
/note="hypothetical protein"

/codon_start=1
/protein_id="BAAB7833.1"

/db_xref="GI:6498430"

/translation="MKLKVDPHPRSHHVFYLAISLTSLIYKTLIYKTLAVPE
KNGTSSLTGSPVSKDSRRNRPGVNLADH"

/complement(join(63589..63842,63844..64145))

/note="hypothetical protein"

/codon_start=1
/protein_id="BAAB7834.1"

/db_xref="GI:6498431"

/translation="MRRSRARPTDLRAGRATAPALRRGRGRSRGRGOEYLGTAISLGGG
EGSTSDMEGRGEGGGAAYRPAQEERRRRRAETKLGSLIYGGRGAPHRQSGHG
EGGRDEEGELMLVGGALRPEEEDEEGHGRRRGRGRGEERRRVCERERER"

Query Match

6.98; Score 75.8; DB 7; Length 142418;

Best Local Similarity 67.3%; Pred. No. 3.6e-08; Mismatches 52; Indels 0; Gaps 0;

Matches 107; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 433 tatctagagctctacaaacaacatccaggaactatacttcgaacttgaaatctg 492
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53805 TACAGAGGCTGTACAAAACAATATTCAGGGCACTATCCAGCTGAGCTCGGAGTCTG 53746
OY 493 aagaatctcatcagcttgatctgtacacaacaatcttaaggagatagttccacccttc 552
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53745 AAGAACCTTATTAGCTTGAGCTGTATACACACATTACTGGAACCATACCGAAGAA 53686
OY 553 ttgggaaattgaagctctgtgctctttttagcgcttaatt 591
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53685 CTTGGCAAGCTGAGTCTCTAGTATCTCTGTGAGTAAGT 53647

Search completed: June 24, 2000, 00:15:32
Job time: 45887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:52:13 ; Search time 446.21 Seconds
(without alignments)
620.139 Million cell updates/sec

Title: US-09-180-798-22

Perfect score: 1106
Sequence: 1 tcgacccgcgcgtccctcca.....aaaaaaaaaaaaaaaa 1106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	1106	100.0	1106	1	V06586	Arabidopsis thaliana
2	924	83.5	981	1	V06587	Arabidopsis thaliana
3	912.4	82.5	1063	1	V06590	Arabidopsis thaliana
4	714.2	64.6	788	1	V06588	Arabidopsis thaliana
5	330	29.8	894	1	V06589	Arabidopsis thaliana
6	191	17.3	2089	1	V06591	Arabidopsis thaliana
7	84.2	7.6	1814	1	V06571	Daucus carota SERK
8	74	6.7	6695	1	V06570	Daucus carota SERK
9	72.4	6.5	4081	1	V06585	Arabidopsis thaliana
10	70.8	6.4	3176	1	T62124	Arabidopsis thaliana
11	64	5.8	3573	1	T06307	Partial tomato pat
12	64	5.8	6471	1	T06306	Tomato pathogen re
13	61.2	5.5	3979	1	V14518	CF-5 pathogen resi
14	61.2	5.5	3979	1	V14519	CF-5 pathogen resi
15	61.2	5.5	4123	1	V14523	CF-5 pathogen resi
16	61	5.5	3293	1	X23532	Tomato Xa21 clone
17	58.6	5.3	6256	1	T31329	Rice Xa21 disease
18	58.6	5.3	7204	1	X23527	O. longistaminata
19	57.6	5.2	3045	1	X23531	Maize Xa21 gene DT
20	57.6	5.2	3921	1	T31300	Rice Xa21 disease
21	57.4	5.2	5940	1	X23520	O. longistaminata
22	57.4	5.2	9424	1	X23525	O. sativa Xa21 gen
23	57.4	5.2	13340	1	X23522	O. longistaminata
24	57.4	5.2	19639	1	X23524	O. longistaminata
25	56.4	5.1	1554	1	T31307	Tomato RRR gene cl
26	55.6	5.0	3541	1	V14522	CF-5 pathogen resi
27	54.8	5.0	8416	1	X23523	O. longistaminata
28	54.4	4.9	2075	1	T49435	Tomato polygalactu
29	53	4.8	1108	1	T49434	Pear polygalactu
30	51	4.6	4104	1	X07356	Arabidopsis sterol
31	50	4.5	2192	1	X23530	Maize Xa21 gene DT
32	49	4.4	5733	1	X00477	Arabidopsis thaliana
33	48.2	4.4	3842	1	X23533	Tomato Xa21 clone
34	46.2	4.2	3050	1	T06309	Tomato pathogen re

35	46.2	4.2	3089	1	T47877	Tomato pathogen re
36	41	3.7	1957	1	Q05644	Clone hns16 encodi
37	40.8	3.7	792	1	Q42596	Sequence of varian
38	40	3.6	5852	1	Q11710	Dictyostelium plas
39	39.6	3.6	19124	1	T72882	Plasmodium var-7 g
40	39.4	3.6	4590	1	N60472	Sequence encoding
41	39	3.5	2880	1	Q21450	Tomato Cf-9 cDNA.
42	39	3.5	2880	1	T06304	Tomato pathogen re
43	39	3.5	3905	1	Q91449	Tomato Cf-9 gene.
44	39	3.5	3905	1	T06303	Tomato pathogen re
45	38.6	3.5	1276	1	T42071	Human nuclear prol

ALIGNMENTS

RESULT	1	
ID	V06586	standard; cDNA to mRNA; 1106 BP.
AC	V06586:	
DT	03-AUG-1998	(first entry)
DE	Arabidopsis thaliana SERK LRR homologous EST clone.	
KW	receptor kinase; apomictis; apomictic; seeds; production; embryos;	
OS	plant breeding; leucine-rich repeat; ss.	
FW	Arabidopsis thaliana.	
FT	Key	Location/Qualifiers
FT	CDS	142..798
FT		/tag=
FT		/note="Shows high homology to SERK"
PN	W09743427-A1.	
PD	20-NOV-1997.	
PF	13-MAY-1997; E02443.	
PR	14-MAY-1996; GE-010044.	
PA	(NOVS) NOVARTIS AG.	
PI	De Vries SC, Hecht VFC, Schmidt EDL, Van Holst GJ.	
DR	WPI; 98-086529/08.	
DR	P-SDB; W47018.	
PT	Production of apomictic seeds - useful in plant breeding	
PS	Claim 28; Pages 71-73; 123pp; English.	
CC	The sequence is that of an EST clone showing high homology to	
CC	SERK LRR (leucine-rich repeat) sequences.	
SQ	Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;	
Query Match 100.0%; Score 1106; DB 1; Length 1106;		
Best local Similarity 100.0%; Pred. No. 8.9e-212;		
Matches 1106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	tcgacccgcgcgtcccaacttaataaagggaaggaacgaactgaactaatttgcct 60
DB	1	TCGACCCGCGCGTCCCTCACTCAATAAAGGGAACCAACGTAACCTTAATTGCT 60
QY	61	ttctcctcttggcagaataatttccttactctcaaatctcttcgaattccctc 120
DB	61	TTCTCCTCTTGTTCAGAAATTTTCCTTACTCAAAATTCCTTTCGATTCCCTCT 120
QY	121	cttaaacctcgaactacatcagcgtctcgaacttcggttggaactcttcgcagc 180
DB	121	CTTAACCTCGAAGCTCATGCGCTCGAAGCTATGAGTGGAGCTTTCGCACT 180
QY	181	tcgttaaccctaacttaacttgatcttaccacgctgctgaagcaaacctcgaagatgct 240
DB	181	TCGTTACCTTAACCTTAGCTTGATTACCTGTCGGAAGCAAACTCGAAGGAGATGCT 240
QY	241	ctttagcctcttgcggcgaggttggaagctcaagacatgtctctcaagcttgatcca 300
DB	241	CTTACGCTCTTGCGCGGAGTTTGACAGATCCAGACCATGCTCTCAAGCTGGGATCCA 300
QY	301	actcttgtaactcttgtaactcgttgatccatgtaactcgttaacgaagacaacgcgtcact 360
DB	301	ACTCTTGTAACTCTGTAACTCGTTCATGTCATGTCATGTAACCAAGAACGGGTACT 360
QY	361	cgtgtgatttggaattcaaacctctctggaacatcttgcgctgagcttggaagctt 420

```

|||||
Db CGTGTGATTTGGAAATTCAAACCTCTCTGGACATCTTGCCGCTGAGCTTGGAAAGCT 420
QY 361
421 gaacatttcagatcttagagctctcaaaaacacatccaaagacatatactctccaa 480
Db 421
421 GAAACATTTAGAGATCTAGAGCTCTACAAAACACATCAAGAAAGAACTATACCTCCGA 480
QY 481
481 ctctgaatctgaagaatctcatcagctctgagctctgtacaaacaaatcttaacaggata 540
Db 481
481 CTTGGAATATCTGAAGAATCTCATCAGCTTGATCTGTACAAACAACTATTACAGGATA 540
QY 541
541 gtcccaactctctctgggaaaaattgaaagtctctggtctttttacggcttaatacaacgga 600
Db 541
541 GTTCCACCTTTCTTGGGAAAATTTGAAGTCTCTGCTTTTACGGCTTATATGACAAACGA 600
QY 601
601 ttgacggctcgaatctccttaagagcactcagcgcaatcccaagccttaaaattgttagctc 660
Db 601
601 TTGACCGGTCCATCTCTAGAGCACTCACGGCAATCCCAAGCCTTTAAATTTGTGACGTC 660
QY 661
661 tcaagcaatgatttctgtgggacaatcccaacaaacggaaccccttgctcaatccctta 720
Db 661
661 TCAAGCATATATTGTGTGGGACAAATCCCAACAAAGGACCCCTTGCTCACAATCTCTTA 720
QY 721
721 cagaactctgaagaacacccggaagattggaaggacggaattactcgtctctgaagctac 780
Db 721
721 CAGAACTTTAGAACACCCGGAATGGAGGACCGGAAATTAACCTTGCTTTCAGACTAC 780
QY 781
781 gaacactatgacactctgaaacaaactgcaaaaaccttgaataagaagattggaagctgac 840
Db 781
781 GACACTATGACACTGAAACAACTGGCAAAACCTGAAATGAAATGGGGGGTGACC 840
QY 841
841 ttgtaagaacactcccaattatcaaatatcacatctattatgaataagatatata 900
Db 841
841 TTGTAAGAACAACCTCACCACTTATCAAAATATCAATCTATTATGTAATGAATATATA 900
QY 901
901 ttagtaaaaaaacaataaatagaagaatcgaatcggtaaatcatccttggtctcaattgag 960
Db 901
901 TTAGTAAAAAACAATAAATGAAGATCGAATCGGAATATCATCTGGCTCAATTTGAG 960
QY 961
961 aactctgaagctctgtatgtaaaatctcctaattcgatcttcgtctctatgtctcgt 1020
Db 961
961 AACTTCAGAGTCTGTATGTAATTAATTTCTAATTCGAATTCGATTCCTTAATGTTCCGT 1020
QY 1021
1021 tctggaactctgaagaatcaatctgattgattgattgattgattgattgattgattgatt 1080
Db 1021
1021 TGtGGATTTCTGAGAAATACATTTGTATGTATGTATGTATGTATGTATGTATGTATGTAT 1080
QY 1081
1081 gcaaaaaaacaataaataaataaataa 1106
Db 1081
1081 GCAAAAAAACAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1106

```

```

PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 75-77; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SRK LRR (leucine-rich repeat) sequences. 180 G; 279 T;
SQ Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 83.5%; Score 924; DB 1; Length 981;
Best Local Similarity 98.8%; Pred. No. 1.3e-225;
Matches 952; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 51 taatttgccttctctctctcttctgttcgaaaaatttcccttctctctcaattcccttttcg 110
Db 13 TTAGTTTGCTTTTCTCTCTTTGTTGCAAAAATTTTCCCTTTACTCTCAATTCCTTTTCG 72
QY 111 attccctccttaaacctccgaagaactcacatgagctctcgaactatcgtctggagact 170
Db 73 ATTTCCTCTCTTAACCTCCGAAGCTCAGATGGGCTTCGAAACTATCGGTGGAGCT 132
QY 171 ctctgagctctgttaaccttaaccttaaccttgattcaacctggtctcgaagaactcga 230
Db 133 CTTCGACGCTTCGTAAACCTTAACCTTAACCTTGATTCACCTGGTCGAAGCAAACTCCGA 192
QY 231 agagagatgctctctacgctctctcggagattgacagatccagaccatgctccagag 290
Db 193 AGGAGATGCTCTCTACGCTCTTGCCGGAGTTGACAGATCCAGACATGCTCTCCAGAG 252
QY 291 ctggatcccaactctgtttaaactctgtacactggtctcaatgctcactgtaacaaagaca 350
Db 253 CTGGGATCCAACTCTGTATATCTTGATACCTTGATCCATGCTGATCCACTGTAAACAGACA 312
QY 351 ccggtctactcgtgtgatttgggaaattcaaacctctcttgacatcttgctgctgagct 410
Db 313 CCGGTACTCGTGTGATTTGGAAATTCAAACCTCTCTGGACATCTTGCCCTGAGCT 372
QY 411 tgggaagcttgaacatttacagatctagaagctctcaaaaacaaacatccaaagaaact 470
Db 373 TGGGAACCTTGAACATTTACAGTATCTAGAGCTCTCAAAAACAACTCAAGAAACTAT 432
QY 471 acctctcgaacttggaaatctgaaatctcatcagacttggatctgtataaacaacatct 530
Db 433 ACCTTCGGAACCTTGAATATGAAATCTCACTACCTTGAGATCTGTAAACAACTAT 492
QY 531 tacagagatagttcccaactctcttggaataattgaagctctctgtcttttcaagcttaa 590
Db 493 TACAGGATAGTTCACCTCTTTGGGAAATGGAAGTCTCTGCTTTTAAAGGCTTA 552
QY 591 tgaacacgatttgacgggtccaat-ectagagcactcaeggcgaatcccaagccttaag 649
Db 553 TGACAAACGATTTGACCGGTCCAATCCCTAGAGCACTCACGGCAATCCCAAGCC-TTAAAG 611
QY 650 ttgttgagctctcaagaatgatttgtgtggacaatcccaacaaacggaaccccttgctc 709
Db 612 TTGTTGACGCTCAAGCAATGATTGTGTGGAACAATCCCAACAACGGAACCTTTGCTC 671
QY 710 acattccttcaagaactctgagaacaacccggaattggaaggacccggaattactcgtc 769
Db 672 ACATTCCTTTACAGAACTTGAGAACCAACCGAGATTGGAGGACCGGAATTCTCTGGTCT 731
QY 770 ttggaagctcgaactgaacttgcaactggaacaaactggcaaacctggaaatgaagaact 829
Db 732 TTGCAAGCTACGACACTTACTGCACTGGAACAACTGGCAAAACCTGGAATAAGAAATAAT 791
QY 830 ggggggtgaccttgtaagaacaacttcaacacttatacaatatacatatgattgatt 889
Db 792 GGGGGGTGACCTTGTAAGAACACTTACCACTTAAATATCAATATCAATATCAATAT 851
QY 890 aagtatatatgtagtaaaaaaacaataaataagaatctgattgattatcatctgtg 949
Db 852 AAGTATATATGTGTGTAATAAACAATAAATGAAGATCGGTATATCAATCACTGG 911
QY 950 tctcaattgagaactcgaagctctgtatgtaaatcttaaatgagcatttgccttaccg 1009

```

```
Db 912 TCCTCAATGGAACCTCGAGCTCTGATGTAAATTTCTAAATGCGATTTCGCCTAAT 971
QY 1010 taat 1013
Db 972 TACT 975

RESULT 3
ID V06590 standard; cDNA to mRNA; 1063 BP.
AC V06590;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
DE receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 106..762
FT /note= "shows high homology to SERK"
FN W09743427-A1.
PD 20-NOV-1997.
PR 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: W47022.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 86-88; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;
```

```
Query Match 82.5%; Score 912.4; DB 1; Length 1063;
Best Local Similarity 95.0%; Pred. No. 1.2e-222;
Matches 1011; Conservative 0; Mismatches 36; Indels 17; Gaps 6;
```

```
Db 483 CAACACATCTTACAGGAGATGTCCTTCTTG6AAAATTGAAGTCTGGTCTT 542
QY 579 ttacagcttaatgacaacgattgaccggtccaat-cciaagactcacggaatccc 637
Db 543 TTTCAGGCTTAATGACACGAGATTGACGGGCGCATCCCTTAGACATCTGCAATCCC 602
QY 638 aagccttaagttgttcagctcacaagatgattgtgtggaacaatcccaacaag 697
Db 603 AAGCC-TTAAAGTTGTGATGTCCTCAAGCAATGATTGTGGAACAATCCCAACAACG 661
QY 698 gacccttgctacattcccttiacagaacttgagaacaacccgagattgagagccgg 757
Db 662 GACCTTTGCTCATTCTTTCAGAACTTGAACAACCCGAGTGGAGGACCGG 721
QY 758 aattactcgtcttgaagctacgacactaacctgacactgaaacaactggaacactga 817
Db 722 AATTACGCGCTTGCAGAGCTACGACACTAATGCAACGAAAAAATTGGCAAAACCTGA 781
QY 818 aatgaagaattgggggggagaccttgaaagacacttcacacattatcaatatcat 877
Db 782 AATGAAGAATTTGGGGGAGACCTTGTAAGAACACTTACACCTTTATCAATATCATCAT 841
QY 878 ctatcgtataatagatataatagtagtaaaacaaaaaaatgaagaatcgaatcgt 937
Db 842 CTACTATGTAATAAGTATATATATGATGCCAA--AAAAAATGAAGAATCGAATCAGT 899
QY 938 aatacatcgtctcgaatgaacttcgagtc--tgatgtaaatattctaactcg 995
Db 900 AATATCATCTGGTCTCAATGAGAACTTGAAGCTGTGTATGTAATAATTCATATCG 959
QY 996 atttcgctaacgtaaatgctgtgtgtggaattcgtgaagaatgaacattgattgatat 1055
Db 960 ACTTTCGCGTACGTATATGTTGGTGTGGGATTTCGAAGAATTAATGATGATAT 1019
QY 1056 ggtatcaagttgttcgctgtctcgaacaaaaaaataaaaaa 1099
Db 1020 GGTATCAAGTTGTTGCTTGTCTGCAAAAAAATAAAAAA 1063

RESULT 4
ID V06588 standard; cDNA to mRNA; 788 BP.
AC V06588;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
DE receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 2..664
FT /tag= a
FT /note= "shows high homology to SERK"
FN W09743427-A1.
PD 20-NOV-1997.
PR 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: W47020.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 79-80; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;
```

```
Query Match 64.6%; Score 714.2; DB 1; Length 788;
Best Local Similarity 96.2%; Pred. No. 2.2e-172;
Matches 753; Conservative 0; Mismatches 28; Indels 2; Gaps 2;
```

```

Db      5  CCCACGCTCCGGCAACTATCGGTGGAGCTTCCGACGCTTGCTTAATCTTAACCTTA 64
Qy      199  gcttgatcaccttgctggaagcaactccgaagagatgctctcactcgtcttcgcgg 258
        |||||
Db      65  GCTTGATTAACCTGCTGGAGCAAACTCCGAGAGATGCTCTTAAAGCTTCCTCCGCGG 124
Qy      239  agttgaagatccagacatgctcctcagaagctgggaatccaactctgttaacctctg 318
        |||||
Db      125  AGTTTAAAGATCCGACATGTTCTCCAGAGCTGGGATCCAACTCTGTAAATCCCTGT 184
Qy      319  acctgtccatgctacatgtaacaaagacacacgcgcactcgtgtggaattggaaat 378
        |||||
Db      185  ACCTGCTTCATGTCATCTGTATACCAACACACCCGCTCATCTGCTGAGATTTGGGAAT 244
Qy      379  tcaaacctctctggaacatctgagcctgagcttggaagcttgaacattacaglatc 438
        |||||
Db      245  TCAAACTCTCTGGACATCTGCGCTGAGCTTGGGAACTTGAACATTTACATATCTA 304
Qy      439  gggctctcaaaaaaataatccaaagaaacttaaccttcgaacttggaaatcgaagaat 498
        |||||
Db      305  GAGCTCTACAAAACACATCCAAAGAACTATACCTTCCGAACTTGGAAATCTGAAGAT 364
Qy      499  ctcatcagcttgatctgtacacaaacatcttaagagatagttcccaattcttgga 558
        |||||
Db      365  CTCATCAGCTTGATCTGTACAAACAATCTTACAGGAGATAGTTCCACATTTCTTGGGA 424
Qy      559  aaatgaagctctcgtgctcttttaacggttaatgaacacgattgacgcgtccaat 617
        |||||
Db      425  AAATTGAAGTCTCTGCTTCTTTTACGGCTTAATGACAAACGATGACGGGCGCAATCCT 484
Qy      618  agagacatccagcagccccaagcctttaaagctgtgacgtctcgaacaaatgattgt 677
        |||||
Db      485  AAGAGCTACACGCAATCCCAAGCC-TTAAAGTTGTATGTCACAAACATGATTTGTG 543
Qy      678  tgggacaatcccaacaaacagaccccttgctacatctcccttaacgaacttgaagaa 737
        |||||
Db      544  TGAACAAATCCCAACAAAGGACCTTTGCTCACAATTCCTTACGAACTTTGAGAACAA 603
Qy      738  ccgagagattggagagcggaaatctcgtgtctgcaagctcagacaaacttaactcga 797
        |||||
Db      604  CCCGAGTTGGAGGACCGGAATTAATCTGCGTTCGAAAGCTAACGACATTAATCTGAC 663
Qy      798  aaacaactggaacaccttgaataatgaggaattggggggtgacacttgaagacactc 857
        |||||
Db      664  AAAAATTTGGCAAACTGAATGAAGATGGGGGTGACCTGTGAAGAACTCTCAC 723
Qy      858  cccttataaataatcaacatctatattgtaataagatatatattgtaataaacaanaa 917
        |||||
Db      724  CACTTATCAAAATCACTACTACTATGTAATGAATATATATATAGTCCAAAAAAA 783
Qy      918  aaa 920
        |||
Db      784  AAA 786

```

```

RESULT 5
V06589
ID V06589 standard; cDNA to mRNA; 894 BP.
AC V06589;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 1..678
FT /tag=a
FT /note="shows high homology to SERK"
PN WO97/43427-A1.
PD 20-NOV-1997.
PE 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.

```

```

PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
P-PSDB: W47021.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 83-84; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;

```

```

Query Match 29.8%; Score 330; DB 1; Length 894;
Best local similarity 71.9%; Pred. No. 7.5e-75;
Matches 532; Conservative 0; Mismatches 115; Indels 93; Gaps 4;

```

```

Qy      225  ctccgaagagatgctctcactacgtcttcgcgcgaagttgacagatccagacatgctc 284
        |||||
Db      15  CTCGGAAGGGAGCGCTCTACGCGCTTCGCGGAGCTTACAGATCCAGCAATGTGT 74
Qy      285  ccgaagctggatccgaactctgttaatcctgttaacctggttccatgctacactgaaca 344
        |||||
Db      75  TCAGAGTTGGATCCAACTCTTGTATATCTTGTACTGTTTCACTGATATCA 134
Qy      345  agacaacgcgcactcgtgtgatttggaaatcaaacctctctgacatcttgagcc 404
        |||||
Db      135  ACACATCAAGTCACTGCTGATTTGGGAATCAAACTATCTGACATCTAGTACC 194
Qy      405  tgaagcttggaaagcttgaacattacaglatc----- 438
        |||
Db      195  TGAAGTTGGAAAGCTTGAACATTAATCTGTATGAATCATCACTCTTTGCTTT 254
Qy      438  ----- 438
Db      255  TGATTAATCGAAAACATTTACATTAATCAAGTACACAFATAACATTTGCTTGAGTATA 314
Qy      438  ---agagctctacaacaaacacacacaaaggaactacacctcgaacttgaagaatctg 494
        |||||
Db      315  TAGTGAACCTACAAAAGAGATTAAGGAACTTACCTCTGAGCTTGGAAATCTGAA 374
Qy      495  gaatcctacagcttgatctgttacaaacaaacatcttaaggaatgattcccaactctt 554
        |||||
Db      375  GAGCTTAATCAAGTTGGATCTGTACAAACAATCTCACCGGAAATCCCATCTCTTT 434
Qy      555  gggaaatggaagctctcgtcttcttcaacgcttaatgacacacgaattgacgcgtcc 613
        |||||
Db      435  GGGAAATTTGAAGTCACTGTTTCTTTCGCGCTTAACGAAACCGATGACCGGCTCAT 494
Qy      614  tccctaaagactcagcgaatcccaagcctttaaagttgttgaagcttcaagcaatgatt 673
        |||||
Db      495  TCCTAGAGAACTCAGATTATTTCAAGCC-TTAAAGTTGTGTATGTCTCAGGGAATGATT 553
Qy      674  tgtgtggacaatcccaacaaacggaaccccttgcacatctccttaacagaacttgaaga 733
        |||||
Db      554  TGTGTGAACAAATTCAGTAAAGAGGACCTTTGAACACATTCCTAATGCAAACTTGAGA 613
Qy      734  acaacccgagattggagggagcgaattactcgtcttgcaagctcagacactaactgca 793
        |||||
Db      614  ACAACCTGAGATTGGAGGAGCAACGAACTAGGCTTTCGCGCTATACGAACTTCA 673
Qy      794  cctggaacacatggaacaccttgaataatgagaatttgggggtgacacttgaagaagac 853
        |||||
Db      674  CTTAAAAAGAGATTGAAGAACTATTAAGAAAGAA-TGTTAGTGACTTGTAAAGAACTCT 732
Qy      854  tcaccacttatcaaatc 873
        |||
Db      733  GTACCAAGTGTGTAAATC 752

```

```

RESULT 6
V06591
ID V06591 standard; cDNA to mRNA; 2089 BP.
AC V06591;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK gene.

```

KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
 KW plant breeding; ds.
 OS Arabidopsis thaliana.
 FH Key location/Qualifiers
 FT CDS 195..2072
 FT /tag- a
 FT /product- SERK protein
 PN M09743427-A1.
 PD 20-NOV-1997.
 PE 13-MAY-1997; E02443.
 PR 14-MAY-1996; GB-010044.
 PA (NOVS) NOVARTIS AG.
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
 DR WPI: 98-086529/08.
 DR P-PSDB; W47023.
 PT Production of apomictic seeds - useful in plant breeding
 PS Claim 27; Pages 91-95; 123pp; English.
 CC The sequence is that encoding SERK, a putative receptor kinase.
 CC It may be used as part of a method of producing apomictic seeds
 CC comprising: (a) transforming plant material with a nucleotide
 CC sequence encoding a protein which in active form in a cell or
 CC cell membrane renders the cell embryogenic; (b) regenerating
 CC the transformed material into plants or carpel-containing
 CC plant parts; and (c) expressing the sequence in the vicinity
 CC of the embryo sac. The apomictic seeds and embryos thus produced
 CC can be developed into plant progeny. This is useful in plant
 CC breeding programs. Controllable and reproducible apomixis provides
 CC many advantages in plant improvement and cultivar development in
 CC the case that sexual plants are available as crosses with the
 CC apomictic plant. Apomixis provides for true-breeding, seed
 CC propagated hybrids and could shorten and simplify the breeding
 CC process so that selfing and progeny testing to produce and/or
 CC stabilise a desirable gene combination could be eliminated.
 CC Apomixis allows plant breeders to develop cultivars with
 CC specific stable traits for such characteristics as height,
 CC seed and forage quality and maturity.
 SQ Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;
 CC

Query Match 17.3%; Score 191; DB 1; Length 2089;
 Best Local Similarity 60.5%; Pred. No. 1,9e-39;
 Matches 314; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 220 gcaaacctcgaagagatgctctctacgcttcgcggaggttgacagatccaagacat 279
 DB 270 GCTATTGTGAAGAGTGAGCTTTCGATCTTGAAGGCTTACTGATCCAAACAAAT 329
 QY 280 gtccctccagagctggtatccaactctgttaactctgttactggttccatgctcgt 339
 DB 330 GTCTTGCAAGCTGGGATCTACGCTAGTGAATCTTGCAATGTTCCATGTCATCTGC 389
 QY 340 aaccagaacaacccgctacactcgltgatttggaatcaaacctctgacactct 399
 DB 390 AACACAGAGAACAGTGTATAGATGATTGGGAAATGAGATGATTCGGCAATTGA 449
 QY 400 ggcgctgagctgggaagcttgaacattacagtatctagagctctcaaaaacacatc 459
 DB 450 GTTCCAGAGACTTGCTGCTCAAGAAATTTGCAATTTGGAGCTTACAGTAACATA 509
 QY 460 caaggaactatatacctcgaacttgaatcgaagaaatccatcaactctgacactctac 519
 DB 510 ACTGGCCGATTCCTAGTAATCTTGAATCTGACAACTTAGAGATTGAGATCTTAC 569
 QY 520 aacaacaatctacaggaagtagttccacattcttgggaaaattggaagctcgtgcttt 579
 DB 570 TTAACAGCTTCTCCGCTCTATTCGGAATCATCTGGGAAAGCTTTCAAAAGCTAGATT 629
 QY 580 ttacggcttaatgacaacagatgacggcttgcacatccatagagcactcaagcagca 639
 DB 630 CTCGCGCTTAACAACAACAGCTCCTACGTGGTCAATCTCTATGTCATGACCAATATTA 689
 QY 640 gaccttaagtggtgagctctcaagcaatgattgtgtggacaatcccaacaacgga 699

DB 690 ACCCTTAACCTGTTAGTATCATCAATAACAGACTCTCTGCTCAGTTCTGACATGCG 749
 QY 700 cccctgtcacatctctcttacagaaacttgagaaacac 738
 DB 750 TCCCTTCACCTCTCAACACCATCAGTTTGGTAATTAAC 788

RESULT 7
 ID V06571 standard; cDNA; 1814 BP.
 AC V06571;
 DT 03-AUG-1998 (first entry)
 DE Daucus carota SERK gene.
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
 OS Daucus carota.
 FH Key location/Qualifiers
 FT CDS 94..1755
 FT /tag- a
 FT /product- SERK protein
 PN M09743427-A1.
 PD 20-NOV-1997.
 PE 13-MAY-1997; E02443.
 PR 14-MAY-1996; GB-010044.
 PA (NOVS) NOVARTIS AG.
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
 DR WPI: 98-086529/08.
 DR P-PSDB; W47013.
 PT Production of apomictic seeds - useful in plant breeding
 PS Claim 21; Pages 47-51; 123pp; English.
 CC The sequence is that encoding SERK, a putative receptor kinase.
 CC It may be used as part of a method of producing apomictic seeds
 CC comprising: (a) transforming plant material with a nucleotide
 CC sequence encoding a protein which in active form in a cell or
 CC cell membrane renders the cell embryogenic; (b) regenerating
 CC the transformed material into plants or carpel-containing
 CC plant parts; and (c) expressing the sequence in the vicinity
 CC of the embryo sac. The apomictic seeds and embryos thus produced
 CC can be developed into plant progeny. This is useful in plant
 CC breeding programs. Controllable and reproducible apomixis provides
 CC many advantages in plant improvement and cultivar development in
 CC the case that sexual plants are available as crosses with the
 CC apomictic plant. Apomixis provides for true-breeding, seed
 CC propagated hybrids and could shorten and simplify the breeding
 CC process so that selfing and progeny testing to produce and/or
 CC stabilise a desirable gene combination could be eliminated.
 CC Apomixis allows plant breeders to develop cultivars with
 CC specific stable traits for such characteristics as height,
 CC seed and forage quality and maturity.
 SQ Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;
 CC

Query Match 7.6%; Score 84.2; DB 1; Length 1814;
 Best Local Similarity 52.1%; Pred. No. 2,3e-12;
 Matches 188; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 339 tgcgcctgagcttgggaagcttgaacattacaglatctagagcttcaaaaaaacat 458
 DB 138 TGAATCTTACCTTGACAAATATGGGTTCTATGACATTTGAGCTTACAGCAATATACAT 197
 QY 459 ccaaggaactatactctccgacttgaactcgaagaactcataagcttgatctgta 518
 DB 198 AAGTGGACCAATTCCTAGTATCTTGGAAATCTGACAAATTTGGTGNCTTGACCTTA 257
 QY 519 caacaacaatctacaggaagtagttccacattcttgggaaaattggaagctcgtgcttt 578
 DB 258 CATGAATAGCTTCTCTGACCTATACCGACACTTAGAGAACTTAACAAGCTTAGATT 317
 QY 579 ttacggcttaatgacaacagatgacggcttgcacatccatagagcactcaagcagca 638
 DB 318 CTTCGCTCTAACAACAACAGCTCCTCTGCTGCAATTCATGTCATGACACTAATATTA 377
 QY 639 agccttlaagtggtgagctctcaagcaatgattgtgtggacaatcccaacaacgga 698

	Matches	142;	Conservative	0;	Mismatches	130;	Indels	0;	Gaps	0;
QY	349	aaccgcgtcactcgtgttgatttggysaaattcaaccctcttgacaactctgagccttgag	408							
Db	995	AAAACCTGGTCGTAGGTGGAATCTGTATTAATATACGCTTTTGCTCTATATTCGGCTCA	1054							
QY	409	cttggagaagcttgacaatttaccgtatctcagagcctcaaaacaacatccagaagact	468							
Db	1095	TGGGGGAAATGTGAACAACCTGCTATAGTTGTATCTTTACAAATACCAAGCTTTGCGCTC	1114							
QY	469	ataactctccgaacttggaaattctgaagaattccatcagcttggatctgttaaaaaaataat	528							
Db	1115	ATTCTCTCTTCATCATGGGGAAATCGAAACAACCTGTCTATAGTTGTATCTTTACAAATACAG	1174							
QY	529	cttacaaggagatagtcocacacttctctggysaaattgaagctctcgtctcttttaccgctt	588							
Db	1175	CTTTCTGGCTATATTCCTGCTTCATTTGGGGAAATCGAACAACCTGTTAGTGTATCTC	1234							
QY	589	aatgacaacacgatgtgaccggtgccaactcctaga	620							
Db	1235	TACATATATCAGCTTTCTGGCTCTATTTCTGTA	1266							

RESULT	12
ID	T06306 standard; DNA; 6471 BP.
AC	T06306;
DR	14-APR-1996 (first entry)
DE	Tomato pathogen resistance gene Cf-2.1.
KW	Pathogen resistant; Cf-2.1; tomato; C.fulvum; Avr 4; Avr 9; fungal;
RN	leaf mould; variegation; ds.
OS	Lycopersicon esculentum.
Key	Location/Qualifiers
FH	5'utr
FT	1..1676
FT	/tag= a
FT	1677..5014
FT	/tag= b
FT	/product= Immature_Cf-2.1-protein
FT	signal_peptide
FT	1677..1745
FT	/tag= c
FT	mat_peptide
FT	1746..5011
FT	/tag= d
FT	5015..6471
FT	/tag= e
PN	
PD	W09J31564-A2.
PF	23-NOV-1995.
PR	11-MAY-1995; G01075.
PR	11-MAY-1994; GB-000394.
PR	23-DEC-1994; WO-G02812.
PR	31-MAR-1995; GB-006658.
PR	07-APR-1995; GB-007232.
PA	(GATS-) GATSBY CHARITABLE FOUND.
PI	Hammond-Kosack KE, Jones DA, Jones JDG;
HP	WI: 96-010949/01.
DR	P-PESDB; R85298.
PT	Increasing plant pathogen resistance by induction of variegation -
PS	Clinm 9; Page 80-83; 13pp; English.
CC	T06306 Is the tomato pathogen resistance gene Cf-2.1. In a new method
CC	this gene is expressed highly in genetic constructs which may be used
CC	to impart a broad range of pathogen resistance, by induction of
CC	variegation, to transgenic plants (or parts or propagules of plants)
CC	containing such constructs. Cf-2.1 imparts resistance to the disease
CC	caused by the leaf mould fungal pathogen Cladosporium fulvum.
CC	C.fulvum contains avirulence (Avr) genes that confer recognition by
CC	plants containing Cf-genes, leading to the activation of host
CC	defence mechanisms to attack the disease.
SQ	Sequence 6471 BP; 2073 A; 1106 C; 1122 G; 2170 T;

Qy	349	aacgsgcgaactgctgtagtttggaagaatcaaacctctcggacaacttcggccctgag	408
	349		
Db	2679	AAAACTGTGTAGGTGAATCTTGTTAATAATACACTTTTCGGCTATTTCTCTTCA	2738
Qy	409	cttgggaagcttgaacatttaccgatctagagctctcaaaaaacaacatccagaagt	468
	409		
Db	2739	TGGGGATCTGAACAACCTGTCTATGTTTATCTTACAAATACACAGCTTCTGGCT	2798
Qy	469	ataccttcggaacttggaaatctgaagaacatcatacagttcgatctgtaacaacaat	528
	469		
Db	2799	ATTCTGCTTATATGGGGAACTCGAACACTTGTCTATGTTGATCTTACAAATATACG	2858
Qy	539	cttaagaaggaatagttcccaactttcttgggaagaattgaagtcctcgtcttlltacgctc	588
	539		
Db	2859	CTTTCGTGCTTATTCCTCGCTCATTTGGGGAACTGAAACAACCTGTCTAGTGTGATCTC	2918
Qy	589	aatgacaacggaattgaccggtcccaatcctaga	620
	589		
Db	2919	TACAAATATACGCTTTCGGGCTCATTTCCGGA	2950

RESULT 13	
VL4518	
ID	VL4518 standard; DNA; 3979 BP.
AC	VL4518.
DT	20-MAY-1998 (first entry)
DE	CF-5 pathogen resistance gene variant #1.
KW	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW	tomato leaf mould; Phytophthora resistance; ss.
OS	Lycopersicon pimpinellifolium.
FH	Key
FT	Location/Qualifiers
FT	653..3560
FT	CDS
PN	MO9743429-AL.
PD	20-NOV-1997.
PF	08-MAY-1997; G01249.
PR	24-SEP-1996; GB-019924.
PR	09-MAY-1996; GB-009681.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI	Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
PI	WPT:98-008895/01.
DR	P-PEDB: W41309.
PT	Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT	for production of transgenic plants resistant to pathogens e.g.
PT	tomato leaf mould C. fulvum in tomatoes
PS	Claim 3: Fig 1a: 75pp: English.
CC	This sequence is an example of the polynucleotide of the invention, and
CC	is able to confer pathogen resistance on a plant. It is one of two tomato
CC	CF-5 gene variants, which offer resistance against the pathogen
CC	Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC	the gene into plant cells and regenerating plants from the cells;
CC	asexually or sexually produced offspring can also be subsequently
CC	produced. Expression of the gene in plant cells can confer pathogen
CC	resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC	Oligonucleotides with sequences complementary to the gene or fragments
CC	of it, are useful in anti-sense techniques to reduce gene expression. The
CC	nucleic acids/polynucleotides are useful as hybridisation probes to
CC	identify other genes/fragments conferring pathogen resistance on plants
CC	e.g. Phytophthora resistance in potatoes. Homologues between Cf-5 and
CC	Cf-9 may be used to identify further resistance genes of this class.
CC	Sequence 3979 BP; 1217 A; 717 C; 698 G; 1347 T;

	Query Match	5.58;	Score 61.2;	DB 1;	Length 3979;	
	Best Local Similarity	48.8%;	Pred. No. 2.1e-06;			
	Matches 155;	Conservative	0;	Mismatches 173;	Indels 0;	Gaps 0.
QY	356	tcactcgtgtagtttgggaatccaacactctctcgacaatttgscgtgagcttggga	415			
Dd	2023	TTACTGACACTATTTTTGGGTAACTACTCTCCTTAATAGCGTCCTAATTCCTGCTTCATTGGGGA	2082			

[illegible]

RESULT	14
VT	V14519
ID	V14519 standard; DNA: 3979 BP.
AC	V14519;
DT	20-MAY-1998 (first entry)
DE	CF-5 pathogen resistance gene variant #2.
KW	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW	tomato leaf mould; Phytophthora resistance; ss.
OS	Lycopersicon pimpinellifolium.
FT	Key
FT	Location/Qualifiers
FT	CDS
FT	653..3560
FT	/*tag= a
PT	/product= CF-5 pathogen resistance gene
PN	W09743429-21.
PD	20-NOV-1997.
PF	08-MAY-1997; G01249.
PR	24-SEP-1996; GB-019924.
PR	09-MAY-1996; GB-009681.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI	Dixon MS, Hatzixanthis K, Jones DA, Jones JDC;
DR	WPI: 98-008895/01.
DR	P-PSDB: W41310.
PT	Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT	for production of transgenic plants resistant to pathogens e.g.
PT	tomato leaf mould C. fulvum in tomatoes
PS	Claim 6: Fig 1b: 75pp: English
CC	This sequence is an example of the polynucleotide of the invention, and
CC	is able to confer pathogen resistance on a plant. It is one of two tomato
CC	Cf-5 gene variants, which offer resistance against the pathogen
CC	Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC	the gene into plant cells and regenerating plants from the cells:
CC	asexually or sexually produced offspring can also be subsequently
CC	produced. Expression of the gene in plant cells can confer pathogen
CC	resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC	Oligonucleotides with sequences complementary to the gene or fragments
CC	of it, are useful in anti-sense techniques to reduce gene expression. The
CC	nucleic acids/polynucleotides are useful as hybridisation probes to
CC	identify other genes/fragments conferring pathogen resistance on plants
CC	e.g. Phytophthora resistance in potatoes. Homologies between Cf-5 and
CC	Cf-9 may be used to identify further resistance genes of this class.
Sequence	3979 BP: 1218 A; 716 C; 698 G; 1347 T;

Query Match	5.58;	Score 61.2;	DB 1;	Length 3979;
Best Local Similarity	48.88;	Pred. No. 2.1e-06;		
Matches 165; Conservative	0;	Mismatches 173;	Indels 0;	Gaps 0;

416 agcttgaaacatcttcagatctcagctcacaanaaacatccacaaggaactatacctt 475

[illegible]

RESULT	15
ID	V14523
AC	V14523 standard; DNA: 4123 BP.
DT	20-MAY-1998 (first entry)
DE	CF-5 pathogen resistance gene clone Hcr2-5D gene.
KM	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
OS	tomato leaf mould; Phytophthora resistance; ss.
TH	Lycopersicon pimpinellifolium.
FT	key
FT	Location/Qualifiers
FT	CDS
FT	/*tag= a
FT	603..3002
PN	/product= CF-5 pathogen resistance gene clone Hcr2-5B
PD	MO9743429-A1.
PD	20-NOV-1997.
PF	08-MAY-1997: G01249.
PR	24-SEP-1996; GE-019924.
PR	09-MAY-1996; GE-009681.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI	Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR	WPI; 98-008895/01.
DR	P-RSDB: M41312.
PT	Tomato gene Cf-5. confers resistance to Cladosporium fulvum - useful
PT	for production of transgenic plants resistant to pathogens e.g.
PT	disclosure of transgenic plants
PS	Tomato leaf mould C. fulvum in tomatoes
PS	Disclosure: Fig 7; 75pp: English.
CC	This sequence is a clone of the polynucleotide of the invention, which
CC	is able to confer pathogen resistance on a plant. The polynucleotide of
CC	the invention is one of two tomato Cf-5 gene variants, which offer
CC	resistance against the pathogen Cladosporium fulvum-5. Transgenic plants
CC	can be produced by incorporating the gene into plant cells and
CC	regenerating plants from the cells; asexually or sexually produced
CC	offspring can also be subsequently produced. Expression of the gene in
CC	plant cells can confer pathogen resistance on a plant e.g. to tomato lea
CC	mould (C. fulvum) in tomatoes. Oligonucleotides with sequences
CC	complementary to the gene or fragments of it, are useful in anti-sense
CC	techniques to reduce gene expression. The nucleic acids/polynucleotides
CC	are useful as hybridisation probes to identify other genes/fragments
CC	conferring pathogen resistance on plants e.g. Phytophthora resistance in
CC	potatoes. Homologies between Cf-5 and Cf-9 may be used to identify
CC	further resistance genes of this class.
Sequence	4123 BP; 1253 A; 742 C; 725 G; 1403 T;

Query Match	5.5%;	Score 61.2;	DB 1;	Length 4123;
Best Local Similarity	48.8%;	Pred. No. 2.1e-06;		
Matches 165; Conservative	0;	Mismatches 173;	Indels 0;	Gaps 0;

[illegible]

```
Db 2227 ATCTAACAACCTTGCTAGGTTGTATCTTACAAATATCAGCTTCTGCTCTATTCCTG 2286
QY 476 CCGAacttgaaatctgaaagaatcctcagcttgatctgtacaacaacaatcttacag 535
Db 2287 CTTCATTTGGCAATATGAGAAATCTGCANAATCTGTCTCAGTGAATGATCATTTG 2346
QY 536 ggaagttcccaacttcttgggaaaaatgaagtcctcgtgcttcttaacggttaagaa 595
Db 2347 GGGAAATTCCTTCATTTGTGTGCAATTTGACATCACTGGAAGTGTGTATATGTCAGAA 2406
QY 596 accgatggaccggtccaatccctagagcaactcaagcgcaatcccaagccttaagttg 655
Db 2407 ACAATTTGAAGGGAAGGAAATGTTGGGTAATATCAAGTGAACCTTCAATTTGT 2466
QY 656 acgtccaagcaatgattgtgtgggacaatcccaaca 693
Db 2467 CGATGTCATCTAATAGTTTCAGAGAGAGAGCTCCCTCA 2504
```

Search completed: June 23, 2000, 22:52:36
Job time: 40552 sec

OTHER INFORMATION: sativa")
US-08-475-891A-1

Query Match 5.3%; Score 58.6; DB 3; Length 6256;
Best Local Similarity 50.8%; Pred. No. 7.6e-07;
Matches 167; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

QY 365 tggatttgggaattcaaacctctctgacatcttgcgctgagcttgggaagcttgaaac 424
DB 2735 tggacttggggaataaatacctgggggagcttcttctaattgcttccaaatttcac 2794
QY 425 att---tacagatctagagctctcaaaaacacatccacaggaactatcccgaaac 481
DB 2795 ctctgcttagtttcttgcacttgaattgaaatgaatcacaggaagattccaaagata 2854
QY 482 ttggaatctgaagaatcctacatcagcttgatctgtacacaacaacatcttcaaggatag 541
DB 2885 ttggcaatcttattggcttacaacatctctatctctgcaacaacaatttcagaggggtcac 2914
QY 542 ttccacattcttgggaataattgaagtcctctgtcttcttaagccttaattgacaacgat 601
DB 2915 ttccatcatcttcttggcgagcgttagaacttaggcattctagtcctacgaaacacact 2974
QY 602 tgacccggtccatccctagagcactcagcgcaatcccaagccttaaggcttgtagcgtct 661
DB 2975 tgagcgggtccatccctagagcactcagcgcaatcccaagccttaaggcttgtagcgtct 661
QY 662 caagcaatgatttgttgggacaatccca 690
DB 3035 gcaccaacaattcagctgttgatacca 3063

RESULT 2
US-08-567-375-1
; Sequence 1, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567.375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004.645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475.891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373.375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 023070-058930
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice
; OTHER INFORMATION: (Oryza sativa)"
US-08-567-375-1

Query Match 5.3%; Score 58.6; DB 4; Length 6256;
Best Local Similarity 50.8%; Pred. No. 7.6e-07;
Matches 167; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

QY 365 tggatttgggaattcaaacctctctgacatcttgcgctgagcttgggaagcttgaaac 424
DB 2735 tggacttggggaataaatacctgggggagcttcttctaattgcttccaaatttcac 2794
QY 425 att---tacagatctagagctctcaaaaacacatcccaaggaactatcccgaaac 481
DB 2795 ctctgcttagtttcttgcacttgaattgaaatgaatcacaggaagattccaaagata 2854
QY 482 ttggaatctgaagaatcctacatcagcttgatctgtacacaacaacttcaaggatag 541
DB 2885 ttggcaatcttattggcttacaacatctctatctctgcaacaacaatttcagaggggtcac 2914
QY 542 ttccacattcttgggaataattgaagtcctctgtcttcttaagccttaattgacaacgat 601
DB 2915 ttccatcatcttcttggcgagcgttagaacttaggcattctagtcctacgaaacacact 2974
QY 602 tgacccggtccatccctagagcactcagcgcaatcccaagccttaaggcttgtagcgtct 661
DB 2975 tgagcgggtccatccctagagcactcagcgcaatcccaagccttaaggcttgtagcgtct 661
QY 662 caagcaatgatttgttgggacaatccca 690
DB 3035 gcaccaacaattcagctgttgatacca 3063

RESULT 3
US-08-587-680A-1
; Sequence 1, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
; OTHER INFORMATION: sativa)"
US-08-587-680A-1
```

```

Query Match          5.38; Score 58.6; DB 4; Length 6256;
Best Local Similarity 50.88; Pred. No. 7.6e-07;
Matches 167; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

QY 365 ttgatttggaatcaacctctctgacatcttcgagcttggaagcttgaaac 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2735 TGGACTTGGAGAAATAAAGCTGGGGGAGTCTTCCATTCGTTTCCATCTTCCA 2794

QY 425 att---tacagtatagagctctacaacaaacacatccaaggaactatacctcgaaac 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2795 CTTCCTTAGATTCTTCTTCACCTTGATTGAAATAGATCACAGGAGCATTCGAAAGGATA 2854

QY 482 ttggaatctggaagatctcatcagcttgatctgttacaacaaacacatccaaggaagtag 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2855 TTGGCAATCTTATTTGGCTTACAAATCTCTATCTCTGCAACAAATAATTGAGAGGCTCAC 2914

QY 542 ttccacattctctgggaatgaagtcctcgtctcttcaagcgttaatgaacacgcat 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2915 TTCCATCATCTGTTGGGCGAGGCTTAGAAACTTAGCATCTCTAGTGCCTTAGGAAACAACT 2974

QY 602 tgaacggttccatcttagagcactcagcggaatcccaagccttaagtgttgagctc 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2975 TTAGCGGTTCCATCCCATTTGGCCATAGGAAATCTTACTGAACTTATATCTTACTGCTCG 3034

QY 662 caagcaatgatctgtgtggaacaccga 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3035 GCACCAACAATTCAGTGTTGATACCA 3063
```

```

RESULT 4
US-08-567-375-3
; Sequence 3, Application US/08567375
```

```

; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Rorald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1..2676, 3520..3918)
; OTHER INFORMATION: /product= "Xa-21"
US-08-567-375-3
```

```

Query Match          5.28; Score 57.4; DB 4; Length 3921;
Best Local Similarity 48.18; Pred. No. 1.3e-06;
Matches 163; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 382 aacctctggaacctctgagccttgagcttggaagcttgaacattacaagtactagag 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1159 AAGATCAGAGAGAACATTCGAAAGATATGCGAATCTTATTTGGCTTACAAATCTCTAT 1218

QY 442 ctctcaaaaacaacatccaaggaactataccttcggaacttggaatctgaaagatc 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1219 CTCTCAACAACAATTCAGAGGCTCTCTCCATCATCTGTTGGGCGAGGCTTAGAAACTTA 1278

QY 502 atcagcttgatctgttacaacaaacatctacagagatagttccacattcttggaaga 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1279 GGCATTCTACTCGCTACGAGAAACAATCTGAGCGGTTCCATCCCGTTGGCCATAGGAAAT 1338

QY 562 ttgaagctctggtctctttagcgcttaatgaacacgcatgagccggtccaatcctagag 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1339 CTTACTGAACCTTAATCTTACTGCTGGCACCACAAATTCAGTGGTTGGATACCATAC 1398
QY 622 cactacggaatcccaagccttaagttgtagcgtcgaagcaatgattgtgag 681
Db 1399 ACACCTCAAAACCTCACAACCTTGTTCATTTAGCCTTTCACACTAATTAACCTTAGTGT 1458
QY 682 acaatcccaacaagcagcccttgctcaccattcctta 720
Db 1459 CCAATACCCAGTGAATTAATTCATTAATTCACACTATCA 1497

RESULT 5

US-08-587-680A-3
; Sequence 3, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1..2676, 3520..3918)
; OTHER INFORMATION: /product="Xa-21"
US-08-587-680A-3

Query Match 5.2%; Score 57.4; DB 4; Length 3921;

Best Local Similarity 48.1%; Pred. No. 1.3e-06;
Matches 163; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 382 aacctctcgbgacatcttcgacctgaagcttgggaagcttgaacattacagatcagag 441
Db 1159 AAGATCAGCAGGAAGGATCCGAAAGATATTGGCAATCTTATGGCTTACAACTCTAT 1218
QY 442 ctctcaaaaaaataccaaggaactatacccttcggaacttggaaatctgaagaatc 501
Db 1219 CTTGCAACAAATTAATTCAGAGGCTCTTCATCATCATCTTGGGCAAGGCTTAAAACTTA 1278
QY 502 atcagcttgcagctgtaacaacaatcttacaggatagttccacttcttgggaaa 561
Db 1279 GGCAATCTACTGCTGCTAGAAAACAACTTGAGCGGTTGATCCGTTGGCATAGAAAT 1338
QY 562 ttgaagctcgtgcttctttagcgtttaaigaacagcattgacgcgtcaatcctag 621
Db 1339 CTTACTGAACCTTAATATCTTACTGCTGGCACCACAAATTCAGTGGATGCATAC 1398
QY 622 cactacggaatcccaagccttaagttgtagcgtcgaagcaatgattgtgag 681
Db 1399 ACACCTCAAAACCTCACAACCTTGTTCATTTAGCCTTTCACACTAATTAACCTTAGTGT 1458
QY 682 acaatcccaacaagcagcccttgctcaccattcctta 720
Db 1459 CCAATACCCAGTGAATTAATTCATTAATTCACACTATCA 1497

RESULT 6

US-08-475-891A-3
; Sequence 3, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Confering
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5992 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

```

: NAME/REV: CDS
: LOCATION: join(512..3149, 3993..4393)
: OTHER INFORMATION: /product= "RRK-B"
: OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
: OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
: OTHER INFORMATION: sativa)"
: JS-08-4/5-891a-3

```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,163
FILING DATE: 03-MAY-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 2307E-540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 421..1401
US-08-238-163-3
```

```
Query Match 4.9%; Score 54.4; DB 1; Length 2075;
Best Local Similarity 50.3%; Pred. No. 6.7e-06;
Matches 190; Conservative 0; Mismatches 181; Indels 7; Gaps 2;
```

```
QY 242 tctacgctctcgcgaggtttacagatccagaccatgtccctccagagctggatccaa 301
    |||||
DB 513 TCTACAAATAAAGAAAGACTTAGGCAATCCTTACCATTTAGCTTATGGGATCCAAACAC 572
QY 302 cttctgttaactcctgttacctgttccatgctcaactgtaaccagaaccgcgtcacc 361
    |||||
DB 573 AGATTGCTGTACTGTACTGATCAATAAATGTGACCGGAAAC---CAACCGGATTAAGT 628
QY 362 gtctgagtttgggaataaaccctctctgacatcttggccttggccttgggaagcttg 421
    |||||
DB 629 CTCTCACCGCTTCCAGACCAATATCTCCGGCAATTCGGGACCGCTCGGAGACTTC 688
QY 422 aacattacagtaactagagctctacaa---aaacaacatccaggaaactataccttcg 478
    |||||
DB 689 CATATCTCGAAACATTTGAAATTCATCATGTACTAATCTCACCGGAAACATTCACCTG 748
QY 479 aacttgaataatcgaagaactatcatcagcttgcctgtgtacacacaaacttaacagg 538
    |||||
DB 749 CAATTGGGAAGCTCACAAATCTCAAAATGTTAAGCTCAGCTTCACCTTAACAGGTC 808
QY 539 taattccacttctcttgggaataatgaagctcctgtgtctttttagcgcttaataagaa 598
    |||||
DB 809 CGATCCCTGAAATTCCTTAGTCAGCTGAAGATTGACGTTGCTCGAGTTGATTAATTAAC 868
QY 599 gattgaccggtccaatcc 616
    |||||
DB 869 AATTACCGGAACAATCC 886
```

```
RESULT 9
US-08-238-163-1
; Sequence 1, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAYTCH, John M.
; APPLICANT: POWELL, Ann
; APPLICANT: STORZ, Henrik
; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
; TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
```

```
STREET: Stewart Street Tower, One Market Plaza
City: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,163
FILING DATE: 03-MAY-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 2307E-540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1058 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1023
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..1058
OTHER INFORMATION: /standard_name= "Pear PGIP cDNA"
US-08-238-163-1
```

```
Query Match 4.8%; Score 53; DB 1; Length 1058;
Best Local Similarity 51.5%; Pred. No. 1.2e-05;
Matches 122; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
```

```
QY 380 caaacctctctggaacatcttgcgcttgaagcttgggaagcttgaacattacagatctag 439
    |||||
DB 347 CCAATCTCACTGCGCCATCCATCCACCCGCAATGCGCAAGCTCAAGACTCAATCTCA 406
QY 440 agcttacaataaagaactcaagaagaactatccttcggaacttggaaatctgaagatc 499
    |||||
DB 407 GGTCTAGCTGACCAACTCTCAAGGCTGTGTCTGACTTCTTCAAGCACTCAAGAAC 466
QY 500 tcatcagttgatctgtacacacaaacttacaaggatagttccacttcttggaa 559
    |||||
DB 467 TCACATCTCTGACCTCTCTTCAACAACCTCACCGGATGCCATCCGAGCTCGCTTCTG 526
QY 560 aattgaagctctctgtcttctttagcgcttaataagaacagattgaacggtccaatcc 616
    |||||
DB 527 AGCTCCAAACCTCGCGCTCTGTGTAGACCGCAATAAAGCTCACAGGTCATATTTC 583
```

```
RESULT 10
US-08-567-375-15
; Sequence 15, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
```


Db 3662 TTGTAATGCAATCCTTAA 3682

RESULT 12

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463

FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313

FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: PTZ9P-F15

US-08-232-463-14

Query Match 4.4%; Score 48.6; DB 1; Length 7218;

Best Local Similarity 3.5%; Pred. No. 0.00047;

Matches 12; Conservative 196; Mismatches 135; Indels 0; Gaps 0;

Db 48 ccccaattgcttcctcccttgctcagaataattcccttaccctaaattcctt 107

Db 1113 yyy 1172

Qy 108 tgaattccctcttaaacctcgaagatcacatgctcgaacacatcggtgga 167

Db 1173 yyy 1232

Qy 168 gcttcgcagctcgttaacacacacacacacacacacacacacacacac 227

Db 1233 yyy 1292

Qy 228 cgaagagatgctcttaagctctgcgcgaggttgacagatccacacatgctccca 287

Db 1293 yyy 1352

Qy 288 gagctgagatcccaactctgttaactcctgttacctggttccatgtaaccacaga 347

Db 1353 yyy 1412

Qy 348 caacgcgtcacctcgtgtggaattggaattcaacctctc 390

Db 1413 yyyyyyyyyyyyyyyyyyygacacaaattcttctatctct 1455

RESULT 13

5169835-11

PATENT NO. 5169835

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO: 11

LENGTH: 1957

5169835-11

Query Match 3.7%; Score 41; DB 7; Length 1957;

Best Local Similarity 55.2%; Pred. No. 0.033;

Matches 80; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Db 953 caatgagaacttgagagctcgtatgtataaattcctaagcgatttgcctactgtaa 1012

Db 1813 caatggaataagcttcttgtaacaaattggaacattgcttgcctcctaagt 1872

Qy 1013 tgcgtgtggtgagatctcgaagtaacattgtatgtatgtatgtatgtatgtatgt 1072

Db 1873 gccccagaattgggaacattatcagagatcatatgattatgtatataaagtactgt 1932

Qy 1073 cctgtctgcaaaaaaaaaa 1097

Db 1933 cacaagttcaaaaaaaaaa 1957

RESULT 14

US-08-244-646-16

Sequence 16, Application US/08244646

PATENT NO. 5744692

GENERAL INFORMATION:

APPLICANT: Cervone, Felice

APPLICANT: De Lorenzo, Giulia

APPLICANT: Salvi, Giovanni

APPLICANT: Albersheim, Peter

APPLICANT: Davilli, Alan

APPLICANT: Bergmann, Carl

TITLE OF INVENTION: Nucleotide Sequences Coding An

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sally A. Sullivan
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,646
FILING DATE: 06-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM 91A 000915
FILING DATE: 06-DEC-1991

```

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: WO PCT/IT/00158
3 FILING DATE: 04-DEC-1992
4 ATTORNEY/AGENT INFORMATION:
5 NAME: Sullivan, Sally A.
6 REGISTRATION NUMBER: 32,064
7 REFERENCE/DOCKET NUMBER: 19-94
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: (303)499-8080
10 TELEFAX: (303)499-8089
11 INFORMATION FOR SEQ ID NO: 16:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 792 base pairs
14 TYPE: nucleic acid
15 STRANDEDNESS: double
16 TOPOLOGY: linear
17 MOLECULE TYPE: cDNA
18 HYPOTHETICAL: NO
19 ANTI-SENSE: NO
20 ORIGINAL SOURCE:
21 ORGANISM: Phaseolus vulgaris
22 STRAIN: Pinto
23 IMMEDIATE SOURCE:
24 CLONE: pAD-1 (DSM NO. 57446926821)
25 FEATURE:
26 NAME/KEY: CDS
27 LOCATION: 2..685
28 FEATURE:
29 NAME/KEY: 3'UTR
30 LOCATION: 686..792
31
32 US-08-244-646-16

```

```

Query Match      3.7%  Score 40.8; DB 2; Length 792;
Best Local Similarity 48.0%  Pred. No. 0.025;
Matches 117; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 382 aaectctctggacatcttggcccttgagcttggaaacttgaacattacaglatcagag 441
      |||||
Db 5 AACCTCTCTCGGTCCAAATCCCGCCGATCGCTAAACACCAACTCCACTATCTCTAT 64
      |||||

QY 442 ctctacaaaacacatcccaaggaaactatcctccgacttggaaactcgaagaatctc 501
      |||||
Db 65 ATCACCCACACCAATGTCTTCGGCCGAATCCCGATTCTTGTGACAGATCAAAACCTTC 124
      |||||

QY 502 atcagcttggatctgfacacaacaacatcttaccaggaatagttcccaacttcttggaaaa 561
      |||||
Db 125 GTCACTCTGACTTCTCTCTACAAAGCGCCCTCTCGGCACCCATACCTCCCTCATCTCTCT 184
      |||||

QY 562 ttgaagctcttgctctttttaaaggcttaatgaacagatggaccggtcccaatccagag 621
      |||||
Db 185 CTCCCAACACTCTGAGGAATACATTCATCGACGGCAACCAATCTCGCGGCCATCCCGGAC 244
      |||||

QY 622 caact 625
      ||
Db 245 TCCCT 248

RESULT 15
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Meckiewicz & No. 5389526rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia

```

```

1 STATE: PA
2 COUNTRY: USA
3 ZIP: 19103
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: PatentIn Release #1.0, Version #1.25
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/07/867,106
12 FILING DATE: 19920625
13
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: AU PJ 7187
16 APPLICATION NUMBER: PCT/AU90/00530
17 FILING DATE: 02-NOV-1989
18
19 ATTORNEY/AGENT INFORMATION:
20 NAME: Feeney, Joanne Longo
21 REGISTRATION NUMBER: 35,134
22 REFERENCE/DOCKET NUMBER: RICE-0002
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE: 215-568-3100
25 TELEFAX: 215-568-3439
26
27 INFORMATION FOR SEQ ID NO: 2:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 5852 base pairs
30 TYPE: NUCLEIC ACID
31 STRANDEDNESS: single
32 TOPOLOGY: linear
33 MOLECULE TYPE: DNA (genomic)
34 ANTI-SENSE: NO
35
36 FEATURE:
37 NAME/KEY: CDS
38 LOCATION: 2378..5038
39
40 FEATURE:
41 NAME/KEY: CDS
42 LOCATION: 2378..5038
43
44 US-07-867-106-2

```

	Query Match	3.6%	Score 40	DB 1	Length 5852	
	Best Local Similarity	47.3%	Pred. No. 0.1	Mismatches 135	Indels 0	Gaps 0
	Matches 121	Conservative 0				
QY	848	aacactcaccacttatactaatacacatcbatbatgtaataagatatataatgtagtga	907			
Db	5580	AATATTAAATTTTAAATTAATAAATAAAAAAAAAAAAAAAAATTTAAATTTAAAAAAA	5621			
QY	908	aaacaacaasaaatgagaatcgatcgatgatcatcatcgctccaatggagaacttcg	967			
Db	55620	AAAAAAAAAAAAAAAAAATTAATAATGCTAAATTGTGAATTTAAATAAATAAATAATGTA	5561			
QY	968	aggctcgtatgttaaattctctaatsgcgaatttcgcttaactgtaaatgcttgatgtggga	1027			
Db	5560	AAGGGTTTTTTTTTAAATATATGACATTTTATTTTAAATCATTTGACAGAATTAAAA	5501			
QY	1028	tttcgtgaagttaaccttgtagtgtatgtagtaacagttgctgtcctltgctgcaaaa	1087			
Db	5500	AATCTTAAACAATTAACAATATTGATTTTTTTTTTTTTTTTTTTTTTTTAAAAAAT	5441			
QY	1088	aaaaaaaaaaaaaaaaa	1103			
Db	5440	TCAATATATTAATAA	5425			

Search completed: June 23, 2000, 22:44:52
Job time: 40238 sec

REFERENCE
1 (bases 1 to 447)
Arabidopsis.
Newman,T., deBruin,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raiknel,N., Somerville,S., Thomashow,M.,
Retzel,E. and Somerville,C.
TITLE
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL
Plant Physiol. 106, 1241-1255 (1994)
MEDLINE
95148729
COMMENT
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcne@lm.cl.msu.edu
Seq primer: T7 dye primer.

FEATURES
SOURCE
Location/Qualifiers

1..447
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="179A227"
/clone_1lb="Lambda-PRU2"
/note="Vector: lambda Zip-Lox. Site.1: Sal. Site.2: Not.
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. " 1190 dt primed cDNA. 126 t 17 others

BASE COUNT 106 a 115 c 83 g 126 t 17 others
ORIGIN

Query Match 34.2%; Score 378.4; DB 23; Length 447;
Best Local Similarity 93.1%; Pred. No. 2.4e-78;
Matches 418; Conservative 0; Mismatches 28; Indels 3; Gaps 3;
QY 17 tccaatcctaataagggaacacagcgttaacccctaatttgccttctccttcttca 76
Db 1 tccaatcctaataagggaacacagcgttaacccctaatttgccttctccttcttca 59
QY 77 gaaatttcccttaccctcaaatctcttgcatttcctcctcctaaccctcgaaag 136
Db 60 gaaatttcccttaccctcaaatctcttgcatttcctcctcctaaccctcgaaag 119
QY 137 ctcaatgagctcgaacactacgcttgagagctcttcgcagcttcgtttaaccctaact 196
Db 120 ctcaatgagctcgaacactacgcttgagagctcttcgcagcttcgtttaaccctaact 179
QY 197 tagctttattcaactggttcgaagaacactccgaagagatgctcttaacgctcttcgc 256
Db 180 tagctttattcaactggttcgaagaacactccgaagagatgctcttaacgctcttcgc 238
QY 257 ggaagtacagatccagacatgctctcagagctggagatccactctgttaactcctt 316
Db 239 ggaagtacagatccagacatgctctcagagctggagatccactctgttaactcctt 298
QY 317 gtaactggttcacatgctcctgtaacgaacacccgctacactcgtgtgatttggaa 376
Db 299 gtaactggttcacatgctcctgtaacgaacacccgctacactcgtgtgatttggaa 358
QY 377 attcaaacctctctggaactctgc-gcctagcttgggaacttggaacttacaagat 435
Db 359 attcaaacctctctggaactctgc-gcctagcttgggaacttggaacttacaagat 418
QY 436 cttagagctctacaaaacacatccaag 464

Db 419 NTAGAGNTTCAAAAAACAATCCAGCG 447

RESULT 3
R89998 556 bp mRNA EST 30-DEC-1997
LOCUS 16353 Lambda-PRL2 Arabidopsis thaliana cDNA clone 187H577, mRNA
DEFINITION
sequence.
ACCESSION
R89998
VERSION
R89998.1 GI:957538
KEYWORDS
EST.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana

REFERENCE
1 (bases 1 to 556)
Arabidopsis.
Newman,T., deBruin,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raiknel,N., Somerville,S., Thomashow,M.,
Retzel,E. and Somerville,C.
TITLE
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL
Plant Physiol. 106, 1241-1255 (1994)
MEDLINE
95148729

COMMENT
On Apr 14, 1993 this sequence version replaced gi:693023.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcne@lm.cl.msu.edu
Seq primer: T7 dye primer.

FEATURES
SOURCE
Location/Qualifiers

1..556
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="187H577"
/clone_1lb="Lambda-PRU2"
/note="Vector: lambda Zip-Lox. Site.1: Sal. Site.2: Not.
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. " 151 a 118 c 123 g 146 t 18 others

BASE COUNT 151 a 118 c 123 g 146 t 18 others
ORIGIN

Query Match 32.4%; Score 358; DB 23; Length 556;
Best Local Similarity 92.1%; Pred. No. 1.4e-73;
Matches 431; Conservative 0; Mismatches 29; Indels 8; Gaps 5;
QY 385 ctctctggaactcttggtcgtgagcttggaagcttggaacatttaagatctagaagtc 444
Db 1 ctctctggaactcttggtcgtgagcttggaagcttggaacatttaagatctagaagtc 60
QY 445 tacaaaaacaacatccaaggaactataccttcgaacttggaaatctgaagatctcatc 504
Db 61 tacaaaaacaacatccaaggaactataccttcgaacttggaaatctgaagatctcatc 120
QY 505 agcttgagcttgtaacaaacaaatcttaagaggaatgttccacttcttggaataatg 564
Db 121 agcttgagcttgtaacaaacaaatcttaagaggaatgttccacttcttggaataatg 180
QY 565 aagctctggtcttttaagcgttaatgacaacagatgaccggttccaat-cttagaaga 623

```

|||||
Db 181 AAGTCTGTGCTTTTACGGCTTAATGACACCGATTCAGCGGGCCATCCCTAGAGCA 240
QY 624 ctaagggcaatcccaagccttaagtgttgacgtctaaagaatgattgttgagac 683
Db 241 CTCACGTGCAATCCCAAGCC-TTAAAGTGTGATGTCTCAAGCAATGATTTGTGTGNNC 299
QY 684 aatccacaacaacgagaccccttgctcacatcccttaccagaacttgagaacaacccgag 743
Db 300 AATCCCAACAAGAGACTTTGCTGCATCTCTTACAGAACTTGAGNACAACCCGAG 359
QY 744 attgg-agggaccggaattactcgtcttgcaagctaacgacatactgcaactgaaaca 802
Db 360 GTTGGNAGGAGACCGAATTACTCGTCTTCAAGCTNCGCAATACTGAC--TGAAA 416
QY 803 actggaacaactgaaatgaagaattgggggtgacctgtgaagac 850
Db 417 AATTGAAAACCTGAAA--TGAGATTTGGGGGTGNCCTTAAAGCAC 462

```

RESULT 4
AA712221 401 bp mRNA EST 24-DEC-1997
LOCUS 31949 lambda-PRL2 Arabidopsis thaliana cDNA clone 180C9T7, mRNA
DEFINITION
ACCESSION AA712221
VERSION AA712221.1 GI:2722138
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 401)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlroge,J., Raikhel,N., Somerville,S., Thomashow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On Sep 19, 1997 this sequence version replaced gi:1520373.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313cne@lhm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1..401
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="180C9T7"
/clone_id="lambda-PRL2"
/note="Vector: lambda zip-lox. site_1: Sal; site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is PRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA. "

BASE COUNT 103 a 105 c 78 g 99 t 16 others
ORIGIN

Query Match 31.1%; Score 343.6; DB 37; Length 401;
Best Local Similarity 92.1%; Pred. No. 3.1e-70;
Matches 363; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

```

QY 121 cttaacctcgaaagctcacatgagcgtctcgaaactatcggtgagagctcttcgaagt 180
Db 1 CTTAACCCNCGAAGAGCTCATGAGGCTCTCGAAACTATCGGTGGAGAGCTCTTCGAGCT 60
QY 181 tcgttaaccctaactagctcttgattcacctggtcgaaagaaactcgaaagagatgct 240
Db 61 TCGTTAACCCCTAACCTTAGCTTATTCACCCNGTGAAGCAACNCCGAGNANTGCT 120
QY 241 cttaagcctcttcgacggaattgagagatccagacatgctctccagagctgagatcca 300
Db 121 CTTCAGCTCTTCGCGGAGATTGACAGATCCAGACCATGCTCTCCAGAGCTGGATCCA 180
QY 301 acctgttaaccttgcttacgtgttcacatgctacactgtaaccaagaacacgcgctact 360
Db 181 ACTMTGTNNATCCTTGTACNGSTTCATGTNACCTGTNACCAAGACAACCGCTCACT 240
QY 421 gaaacttaagatgactagagctctcaaaaacaacataccaagaactatacttcgaa 480
Db 301 GAACATTACGATCTTAGAGCTCTTACAAAACCAACACATCAAGACTATACCTCCGAA 360
QY 481 ctggaaactgaaagatctcatcagcttgagct 514
Db 361 CTGGAATCTGAAG-ATCTCATCAGNTTGGATC 393

```

RESULT 5
AA394359 534 bp mRNA EST 30-OCT-1997
LOCUS 25942 lambda-PRL2 Arabidopsis thaliana cDNA clone 305G1T7 3, mRNA
DEFINITION
ACCESSION AA394359
VERSION AA394359.1 GI:2047570
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 534)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlroge,J., Raikhel,N., Somerville,S., Thomashow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On May 18, 1995 this sequence version replaced gi:811121.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313cne@lhm.cl.msu.edu
Seq primer: T7.
Location/Qualifiers
1..534
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="305G1T7"
/clone_id="lambda-PRL2"
/note="Vector: lambda zip-lox; site_1: Sal; site_2: Not;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source
Location/Qualifiers
1..534
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="305G1T7"
/clone_id="lambda-PRL2"
/note="Vector: lambda zip-lox; site_1: Sal; site_2: Not;

SOURCE		thale cress.
ORGANISM		Arabidopsis thaliana
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 442)
AUTHORS		Neman,T., deBrujin,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Ratkhe,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.
TITLE		Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cdna clones Plant Physiol. 106, 1241-1255 (1994) 95148729
JOURNAL MEDLINE COMMENT		On Jan 14, 1998 this sequence version replaced g1:1797474. Contact: Thomas Neman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel.: 517-353-0854 Fax: 517-353-9168 Email: 223jstc@ibm.c1.msu.edu The sequence entry for this EST has been reverse complimented and is being submitted in the sense orientation. Seq primer: M13.universal. Location/Qualifiers 1..442 /organism="Arabidopsis thaliana" /strain="var columbia" /db_xref="taxon:3702" /clone="90JXP" /clove_1lb="lambda-PRL2" /note="vector: lambda Zip-lox; site_1: SalI; site_2: NotI; Lambda PRL2 is a CDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and silique). The vector is BRU's lambda Zip-lox. The CDNA inserts were directionally cloned with Sal-I/Nci arms using oligo dt primed cDNA."
FEATURES SOURCE		
BASE COUNT ORIGIN	135 a 82 c 90 g 132 t 3 others	
Query Match	29.6%; Score 337.8; DB 42; Length 442;	
Best Local Similarity	89.4%; Pred. NO. 1.5e+66;	
Matches 397; Conservative 0; Mismatches 40; Indels 7; Gaps 4;		
OY 644	ttaaagtgttgtagctccaagcaatgatatttggtygaacaac-tccaaacaacygacc 702	
Db 1	TAAAGTTNTTNGATTCCTNCACGANCANGTTTTGTGGAACACTCCCAACAAGAAGACTR 60	
OY 703	tttgtcacattcccttaagaactttgagaacaccccgaggattgg--aggagaccgyaat 760	
Db 61	TTGCCTCACATCTCCTTTACAGACTTGTAAGAACAACC CGAGGTGGAGSGACCGGAAT 120	
OY 761	tactcgcgttcgcaaagtacgacctaacgtgacctgaaaacaactgycnaaacctgaaaa 820	
Db 121	TACTCGGCTTWCAGACTTAGACNACNATACTGCCACTGAAAAAATTGCGAAAACCTGGAAA 180	
OY 821	tgaagaattggggggtagacctgtlaagaacacatttaccaacttttaataatcatcata 880	
Db 181	TGAAGAANTGGGGGCTACCTTTAAGAACACTTCCCACCTTATCCAARTTCCACACTTA 240	
OY 881	ttatgtataaagtatatcatatgtagfaanaaaaaaaaaaabaatgaagaatcgtaatgfaat 940	
Db 241	C TAGTGTAATAGTATATATATGTAGTCCAA--AAAAAAAAAGAGAAATCGAATCATGTAAT 298	
OY 941	atcatcgtgtccaattgagaactcgagtc--tgatgtfnaaaattctaattgagat 998	

Db	299	ATCATGTGCTCAATTCAGAACTTTCAGAGCTGCTGATGTAAATAATTTCTAATCTGACT	358
QY	999	ttgccttactgaagatggttcggttttggtgattctggagaagcaaatcttgattggtatg	1058
Db	359	tttgcctactgtaaaatttcgggttggtggatcttcgagaatgtaacatttggatgg	418
QY	1059	atcaagttgtctgccttgctctgc	1082
Db	419	ATCAAGTTGTCTCGCCTGTCTGCG	442
RESULT	8		
LOCUS	234187	336 bp	EST
DEFINITION	ATTC3221 Versailles-VB Arabidopsis thaliana CDNA clone VBVDH02, mRNA sequence.		14-JUN-1994
ACCESSION	234187		
VERSION	234187.1	GI:496544	
KEYWORDS	EST.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta: euphyllophytes; Spermatophyta: Magnoliophyta: eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 336)		
TITLE	CNR5.		
JOURNAL	The Arabidopsis thaliana transcribed genome: the GDR cDNA program Unpublished (1996)		
COMMENT	Contact: Desprez T., Amselem J., Chlapello H., Rouze P., Caboche M., Hofte H.		
FEATURES	INRA Versailles Laboratoire de Biologie Cellulaire Route de Saint-Cyr, 78026 Versailles Cedex, France Email: thierry@versailles.inra.fr.		
source	Location/Qualifiers 1..336 /organism="Arabidopsis thaliana" /strain="ecotype Columbia" /db_xref="taxon:3702" /clone="VBVDH02" /clone_11b="Versailles-VB" /tissue="whole seedlings" /dev_stage="in vitro-grown etiolated seedlings, 5 days old" /note="Vector: pBluescript"		
BASE COUNT	89 a 62 c 65 g 119 t	1 others	
ORIGIN			
Query Match	29.6%; Score 327.2; DB 20; Length 336;		
Best Local Similarity	98.8%; Pred. No. 2e-66; 4; Indels 0; Gaps 0;		
Matches 329; Conservative	0; Mismatches		
QY	671	attgtgtggacaatcccaacaacgacccttgcctacatccttlacagaacttg	730
Db	333	ATTGTGTGTAACAATCCACAAACGACCCTTTCACACANCTTTACAGAACTTG	274
QY	731	agacaaccccgaaattggagggaccggaattacttcggtcttgaagatgaacactaa	790
Db	273	AGAACCAACCCGAATTTGAGGAGCCGGAAATTCCTGCTTCGCAAGCTACGACATACT	214
QY	791	gcacctaaacaactggcaaacctcgaaaatggaagaattgggggtgtaaccttga	850
Db	213	GCACTGAACAACTGCGAAAACCTGGAATAAGAAATGGGGGCTACTTGAAGAC	154
QY	851	acttaccacttatcaaatatcatcatctatgatgtaaatatataatagtaagta	910
Db	153	ACTTCACCACTTATATCAATATCAATCATCTATTATGTATATATGTATGTAGTAA	94
QY	911	acaaaaaaatgaagaatcgaaatcgatataatcctgcttcacattgaaacttg	970
Db	93	ACAAAAAAATGAAGATCGAATCGGTAATATCACTCTGCTCAATTGAACCTTGAGG	34

```

QY 971 tctgtatgtaaaattcttaaatcgatttcgc 1003
Db 33 TCTGTATGTAATAATTCTTAAGTCGATTTCGC 1

RESULT 9
AL100685 360 bp mRNA EST 21-AUG-1998
LOCUS 33806 Lambda-PRL2 Arabidopsis thaliana cDNA clone 90J7XP 3', mRNA
DEFINITION sequence.
ACCESSION AL100685
VERSION AL100685.1 GI:3449395
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsi.
REFERENCE 1 (bases 1 to 360)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On May 5, 1995 this sequence version replaced gi:797715.
CONTACT: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@dm.cl.msu.edu
The sequence entry for this EST has been reverse complemented and
is being submitted in the sense orientation.
Seq primer: M13-21.
FEATURES
SOURCE location/Qualifiers
1..360
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="90J7XP"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo df primed cDNA."
BASE COUNT 116 a 60 c 77 g 105 t 2 others
ORIGIN
Query Match 28.6%; Score 316; DB 42; Length 360;
Best Local Similarity 95.6%; Pred. No. 8,3e-64;
Matches 346; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
QY 723 gaacttgaagaacaaccgagattgagagaccggaattctcggtctgcaagctacga 782
Db 1 GAACCTTAGAACACCCGAGGTTGAGGGACCGAATTACTCGTCTTCCAGCTACGA 60
QY 783 cactactgcacctgaacacctggcaaaacctgaataaattgagggtgacctt 842
Db 61 CACTACTGCACCTGCAAAAAATTGGCAAAACCTGAAATGAAGATTGGGGGTGACCTT 120
QY 843 gtaagaacacttcaccacttatacaatatactattatgtaataagtatataatg 902

```

```

Db 121 GTAGAGACACTTCACACCTTATCAAAATATCACACTCTATGTAATATATATATG 180
QY 903 tagtaaaacaaaanaaatagaagaatcgatcgtaataatcatctgctcctaattgaa 962
Db 181 TAGTCCAA--AAAAAAGAAAGAAATCGAATCAGTAATATATCTGAGCTCAATGAGAA 238
QY 963 cttagagatc--tgatgtaaaattcttaaatcgatttcgactgtaagtgcgt 1020
Db 239 CTTTGAGGTCCTGTGATGTAATAATTTCTAAATGCGACTTTCGCGACTGATGTCGGT 298
QY 1021 tctggattctgagaagatgatacttgatgtagatgataagtgtgcttcctct 1080
Db 299 TGTGGATNCTGAGAAAGTAACATTGTATGTGTATGATGATCAAGTTGTCGCTTGCT 358
QY 1081 gc 1082
Db 359 GC 360

RESULT 10
H37300 353 bp mRNA EST 30-DEC-1997
LOCUS 15429 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179J19T7, mRNA
DEFINITION sequence.
ACCESSION H37300
VERSION H37300.1 GI:906799
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsi.
REFERENCE 1 (bases 1 to 353)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On May 8, 1995 this sequence version replaced gi:801146.
CONTACT: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@dm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
SOURCE location/Qualifiers
1..353
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="179J19T7"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo df primed cDNA."
BASE COUNT 98 a 85 c 67 g 89 t 14 others
ORIGIN

```

Query Match 28.0%; Score 309.6; DB 23; Length 353;
Best Local Similarity 94.4%; Pred. No. 2.6e-62;
Matches 334; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

445 tacaataaacatcccaagaactatactccgaacttggaatctgaagaatctcacc 504
|||||
1 TACAAAACACATCCACAGACTATACCTTCGAACTTGGAAATCTGAAGATCTCACC 60
|||||
QY 505 agcttgatctgtacacaacaactctacaggaatagttccactctcttggaattg 564
|||||
DB 61 AGCTTGCTGTGTACACAACAATCTTACAGGATAGTCCCTTTTGGGAAATTTG 120
|||||
QY 565 aagctctgtcttttagagcttaataagacaaccgattgacgggtccaat-cttagagca 623
|||||
DB 121 AAGTCTGTGCTTTTACGGCTTAATGACACCGATTGACCGGTCCATCCCTAAGCA 180
|||||
QY 624 ctacagcaatcccaagcctttaagtgtgtgacgtctcacaagaatgattgtgtggagc 683
|||||
DB 181 CTCACGGCAATCCCAAG-CATTAAAGTTGTGANGTCTCANGCAATGATTGTGTGAAC 239
|||||
QY 684 aatcccaacaagagacccttctgacatctcttcaagaatttgagaacaaccggag 743
|||||
DB 240 AATCCCAACAACGCGNCCCTTGTCTCAGTCCCTTACAGAACTTGGNAAACCCGNG 299
|||||
QY 744 atggaagagaccggaattactcgtctgtgcaagctacgacactgaactgacctg 797
|||||
DB 300 ATTGAGAGGNCGGANTTACTCGTCTTGCAGATACGGCACTAATGACANGC 353
|||||

RESULT 11
A1100683 364 bp mRNA EST 21-AUG-1998
LOCUS 33804 Lambda-PRL2 Arabidopsis thaliana cDNA clone 187H5XP 3', mRNA
DEFINITION
sequence.
ACCESSION A1100683
VERSION A1100683.1 GI:3449393
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On May 5, 1995 this sequence version replaced g1:797711.
COMMENT
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu
The sequence entry for this EST has been reverse complimented and
is being submitted in the sense orientation.
Seq primer: M13-21.

FEATURES
SOURCE

1. 364
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="187H5XP"
/note="Vector: lambda Z1p-Lox; site 1: Sal; site 2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7

day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda Z1p-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. " others
BASE COUNT 113 a 80 c 70 g 99 t 2

Query Match 27.7%; Score 306.2; DB 42; Length 364;
Best Local Similarity 95.3%; Pred. No. 1.6e-61;
Matches 347; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

544 cccacttcttgaggaaattgaatctctgtctcttttaagggcttaatgaaccaggtt 603
|||||
DB 1 CCCATTTCTTTGGGAAATTAAGTCNT-GTCTTTTACGGCTTAATGACACCGATTG 59
|||||
QY 604 accggtccaatcc-tagaagcactcagcgaatcccaagccttaagtgtgtgacgtc 662
|||||
DB 60 ACGGGCCCAATCCNTAGACACTGACATGCCAAGCC-TTAAATGTGTATGTCTC 118
|||||
QY 663 aagcaatgattgtgtggacaatcccaacaagagacccttgcacatcttaca 722
|||||
DB 119 AAGCAATGATTGTGTGGAACAATCCCAACAGACCTTTCTCACAATCCCTTACA 178
|||||
QY 723 gaacttgagaacaaccggagattggagagcgggaattactcgtgttggaagctcga 782
|||||
DB 179 GAACCTTGGAACAACCCGAGGTGGAGGACCGGAATTAATCTGCTGTGACACTACGA 238
|||||
QY 783 cactactgcacctgaacaactgcaaacctgaaatgaagaattgggggtgacctt 842
|||||
DB 239 CACTACTGCACTGAAATAATGGCAAAACCTGAAATGAAGATTGGGGGGTGACCTT 298
|||||
QY 843 gtaagaacacttaccacttatacaatatacatctatgttaataagatatatg 902
|||||
DB 299 GTAAGACACTTACCACTTATCAATATATCAACATCTACTATGTAATATATATG 358
|||||
QY 903 tagt 906
|||||
DB 359 TAGT 362

RESULT 12
Z34606 337 bp mRNA EST 11-JUL-1994
LOCUS A1133366 Versailles-VB Arabidopsis thaliana cDNA clone VBVDH02,
DEFINITION
mRNA sequence.
ACCESSION Z34606
VERSION Z34606.1 GI:507009
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS CNRS.
TITLE The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
COMMENT Contact: Desprez T., Amselem J., Chiapello H., Rouze P., Caboche
M., Hofte H.
INRA Versailles
Laboratoire de Biologie Cellulaire
Route de Saint-Cyr, 78026 Versailles Cedex, France
Email: therry@versailles.inra.fr.

FEATURES
source

1. 337
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia"
/db_xref="taxon:3702"

/clone="VBVDH02"
/clone_lib="Versailles-VB"
/tissue_type="whole seedlings"
/dev_stage="in vitro-grown etiolated seedlings, 5 days old"
/note="Vector: pBluescript"
BASE COUNT 73 a 98 c 60 g 105 t 1 others
ORIGIN

Query Match 27.6%; Score 305; DB 20; Length 337;
Best Local Similarity 97.6%; Pred. No. 3e-61;
Matches 330; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 55 ttgtcttcctcctctgttcagaaatttcccttactctcaaatcccttgcattt 114
DB 1 ttgtcttcctcctctgttcagaaatttcccttactctcaaatcccttgcattt 60
QY 115 cctctcttaaacctcgaagaatgcacatgagctcgaactatcggtggagctctc 174
DB 61 CCTCTCTTAACCTCCGAAAGCTCAGATGGGCTCTCGAAATATCGGTGGAGCTCTTC 120
QY 175 gcaagctcgttaaacctcgaagaatgcacatgagctcgaactatcggtggagctctc 234
DB 121 GCAAGCTCGTTAACCTCACTAGCTTGATTCACCTGTCGAGCAAACTCCGAA-GA 179
QY 235 gatacctctacagctcttgcgagagttgacagatccagacatgctctcgaagctcg 294
DB 180 ATGCTCTACAGCTCTTCGCGAGATTGACAGATCCAGACATGCTCTCCAGAGCTGG 239
QY 295 gatacctctctgttaacctcgttgcacatgacatgacatgacatgacatgacatgac 354
DB 240 GATCAACACTCTGTATATCTGTACCTGTTCAGATGATGATGATGATGATGATGATGAT 299
QY 355 gtaacctcgtt-ggatttgggaattcaacatcctctcgt 391
DB 300 GTCACTCGTGTGGATTGGGAATTCACACTCTCTG 337

RESULT 13
A1100682 348 bp mRNA EST 21-AUG-1998
LOCUS 33803 Lambda-PRL2 Arabidopsis thaliana cDNA clone 185L1XP 3', mRNA
DEFINITION
sequence.
ACCESSION A1100682
VERSION A1100682.1 GI:3449392
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 348)
Newman,T., deBrujin,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
On May 5, 1995 this sequence version replaced g1:797667.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PR, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 2233tcm@dm.cl.msu.edu
The sequence entry for this EST has been reverse complemented and
is being submitted in the sense orientation.
Seq primer: M13-21.
Location/Qualifiers

FEATURES

source
1. 348
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="185L1XP"
/clone_lib="Lambda-PRL2"
/note="Vector: Lambda ZIP-lox; Site-1: Sal; Site-2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BPL's lambda ZIP-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT 110 a 74 c 66 g 93 t 5 others
ORIGIN

Query Match 26.4%; Score 292; DB 42; Length 348;
Best Local Similarity 95.8%; Pred. No. 3.3e-58;
Matches 320; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 574 gtcttttacggcttaatgacacagatgacggtccatcc-tagaagcactcagcga 632
DB 14 GTCTTTTACGGCTTAATGACACAGATGACGGGCCAATCCATAGACACTGCA 73
QY 633 atcccaagcctttaaagtgttgacgtctcaagaatgattgtgtggcaatcccaac 692
DB 74 ATCCCAAGCC-TTAAAGTGTGATGCTCAAGCAATGATTGTGGAACAAATCCCAAC 132
QY 693 aaacgagaccccttgcacatccttcaagaacttgaagaacacccgagatggagg 752
DB 133 AATCGAACCTTTTGCTACATTCCTTACAGACTTGAAGACAAACCGAGGTGGAGG 192
QY 753 accggaattacgcgttcttgcagactacgactacgactacgactacgactacgactacg 812
DB 193 ACCGGAATTACGTGNTCTGCAAGCTACGACACTGACACTGCAAAAATTTGGCAAA 252
QY 813 cctgaaagaagaatttgggggtgaccttgaagaacacttaccattataatat 872
DB 253 CCGAAAGAAGAANTGGGGGTGANTCTGTGAAGACACTTCACCACTTATCAATAT 312
QY 873 cacatctatcatgaataatataatagtagt 906
DB 313 CACATCTAGATGATATAGTATATATAGTAGT 346

RESULT 14
H37296 336 bp mRNA EST 30-DEC-1997
LOCUS 15425 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179J11T7, mRNA
DEFINITION
sequence.
ACCESSION H37296
VERSION H37296.1 GI:906795
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 336)
Newman,T., deBrujin,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On May 8, 1995 this sequence version replaced g1:801143.
Contact: Thomas Newman

MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PR, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ldm.cl.msu.edu
Seq primer: 77 dye primer.
Location/Qualifiers

FEATURES

source

1. .336
/organism="Arabidopsis thaliana"
/strain="var columbiana"
/db_xref="taxon:3702"
/clone="179117"
/clone.lib="lambda-PRU2"
/note="Vector: lambda ZAP-Lox; Site_1: Sal; Site_2: Not; lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda ZAP-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dr primed cDNA."

BASE COUNT

79 a 86 c 69 g 89 t 13 others

ORIGIN

Query Match 26.4%; Score 291.8; DB 23; Length 336;
Best Local Similarity 92.3%; Pred. No. 3.6e-58;
Matches 310; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

QY 151 cgaactatcgtgtggaagctcttcgacgttcgttaaccttaaccttgattcac 210
DB 1 CGAATCATGTGGTGGAGCTCTCGACGCTCTTAATCCTAACCTTACCTTGATTGATC 60
QY 211 ctggtcgaagcaactcgaagagagatgctctacgcttcgacgagttgacagat 270
DB 61 CTGCTGCAACCAACTCCGAGAGATNGCCTTTAGCGCTTCGCCGAGTTTACAGAT 120
QY 271 ccagacatctcctccagagcttggaatccaaactctgttaactctgtacctggttcac 330
DB 121 CCGGACCATCTTCCAGACTGGATCCAACTGTGAATCCTGTACTCGGNNCCAT 180
QY 331 gtacacctgtacccaagacacgcgtcactcgtgtgatttggaattcaaacctctct 390
DB 181 GTCACTGTACCAAGACAAACGCGTCACTGCTGATTTNAGGATTTAAACCTCTCT 240
QY 391 ggaacatctgcgcctgagcttggaagcttgaaacttacaatctagaagcttacaac 450
DB 241 GGNCACTATTNCGCTTNGCTTTGGAAGCTTGAAACANTTACAGATCTAGAGCNCATCAA 300
QY 451 aacaaatccaagaagacta-taacctccgaacttg 485
DB 301 AACAACTNCAAGGGGCTATTACCTTCGAACTTGG 336

RESULT 15

AM443205

LOCUS 608 bp mRNA EST 17-FEB-2000
DEFINITION EST308135 tomato mixed elicitor, B71 Lycopersicon esculentum cDNA
clone cLET43M10 5', mRNA sequence.

ACCESSION

AM443205

VERSION

AM443205.1

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

Lycopersicon esculentum
tomato.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 608)

AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Viston, T., Holt, I.E.,
Liang, F., Hansen, T.S., Ronning, C.M., Craven, M.B., Bowman, C.L.,
Mierman, M., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
Giovannoni, J.J., and Martin, G.B.
Generation of ESTs from tomato callus (mixed elicitor)
Unpublished (1999)
On Dec 20, 1995 this sequence version replaced g1.1135853.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: df@chelemson.edu
5 prime sequence.

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .608
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET43M10"
/clone.lib="tomato mixed elicitor, B71"
/tissue.type="leaf"
/dev.stage="4-6 week old plants"
/lab.host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLET - inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, Eix,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

BASE COUNT

165 a 126 c 134 g 183 t

ORIGIN

Query Match 25.9%; Score 287; DB 81; Length 608;
Best Local Similarity 68.7%; Pred. No. 4.8e-57;
Matches 395; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 178 gcttcgttaaccccaactagcttgattgattcactcgtggaagcaactccgaagagat 237
DB 33 GGTGTGTGGAGTGTCTTGTGCTGTCGTGCTGTAAAGGGAATTCAGAGGGGAT 92
QY 238 gctctcagctctcgcgcgagtttgacagatccagacatctccccaagagcttgat 297
DB 93 GCTTTGAGACGCCCGCGGAGGATTATCGACCGGTAAGCTGTATACAGAGCTGGAT 152
QY 298 ccaactctgttaactcttgtaacctggttccatgtaactcgttaacaaagacacgcgtc 357
DB 153 CCAATCTGTGTAAACCTTGTACTGTTTCATGTCACTTGCAACGAGATATATCAAGTT 212
QY 358 actcgttggaattggggaattcaaacctctcggacatctcgcctgagcttggaag 417
DB 213 ACTGCTGGGATCTTGGGAACCTAAAGTATCTGTGATTTGTTACTTACAGCTGGAAAG 272
QY 418 ctgaaacttcaagatctagagctctacaacaaacaaatccaagaagataactctcc 477
DB 273 CTTGAACATCTACAGATCTGAGACCTTTACAAAAATATATTCACGACCACTCAAG 332
QY 478 gaacttgaaatctgaaagaaatctcatcagcttgatctgtaacaaacaaatcttaagag 537
DB 333 GAGCTGGTACTTGAAGAGCCTTATTAGTGTGATCTGTACAAACAAATATTTTGGGG 392
QY 538 atagttcccaactcttggaagaaattgaagctctcgtgctttttaagagcttaagcaac 597
DB 393 ACAATTCCTACTTCACTTGGAAACCTGAAAACCTTGTCTTTCGCGTCAATATGATAC 452
QY 598 cgattgacccgttccaatccatctagagcactcaagcaatcccaagcctttaagtgtgac 657
DB 453 AAGCTAACAGGACCAATCCCAAGAACTTACTAGCAATTTCTAGCTGAAAGTTGGGAT 512
QY 658 gtccaagcaatgattgtgtggaacatcccaacaaagaccccttgctacatctct 717

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:15:32 ; Search time 12463.9 Seconds

(without alignments)
-76.566 Million cell updates/sec

Title: US-09-180-798-24

Sequence: 981
1 agtgcgagtaattcagttcgtg.....tcgcctaataattcacaact 981

Scoring table: IDENTITY_MWC
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_cm:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pl4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pl5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	981	100.0	981	5	A67819	A67819 Sequence 24
2	924	94.2	1106	5	A67817	A67817 Sequence 22
3	841.8	85.8	1063	5	A67825	A67825 Sequence 30
4	742	75.6	789	5	A67821	A67821 Sequence 26
5	356.2	36.3	894	5	A67823	A67823 Sequence 28
6	210.2	21.4	2089	5	A67827	A67827 Sequence 32
7	156.6	16.0	936	8	SB062279	U62279 Sorghum b1c
8	101.8	10.4	1755	8	DC093048	U93048 Daucus caro
9	101.8	10.4	1815	5	A67797	A67797 Sequence 2
10	96.6	9.8	4604	7	LEIRGENE	X95269 L.esculentu
11	83.2	8.5	101284	5	ATK009991	AC009991 Arabidops
12	81.4	8.3	81513	50	AC021198	AC021198 Arabidops
13	80.8	8.2	134402	8	ATK011765	AC011765 Arabidops
14	80	8.1	112529	44	AC015446	AC015446 Arabidops
15	79.8	8.1	100887	50	ATK015446	AC015446 Arabidops
16	79.2	8.1	116944	8	AC008017	AC008017 Arabidops
17	78.2	8.0	5033	8	INT77888	U77888 Ipomoea nil
18	77.2	7.9	108355	50	AC005957	AC005957 Arabidops
19	76.2	7.8	84440	7	AB010075	AB010075 Arabidops
20	76.2	7.8	118086	7	ATK6H11	AL021684 Arabidops
21	75.8	7.7	142418	7	AP000815	AP000815 Oryza sat
22	75.4	7.7	133840	8	AC003981	AC003981 Complete
23	75	7.6	5940	8	OLU72727	U72727 Oryza longi
24	74	7.5	6695	5	A67796	A67796 Sequence 1
25	73.8	7.5	2786	7	AB029327	AB029327 Nicotiana
26	73.8	7.5	3176	5	E12705	E12705 Arabidops
27	73.8	7.5	3176	8	ATU47029	E12705 Arabidops
28	73.2	7.5	23143	8	F15P11	U47029 Arabidops
29	72.4	7.4	4081	5	A67815	AF160760 Arabidops
30	72.4	7.4	98471	7	AC012654	A67815 Sequence 20
31	72	7.3	87434	7	AB010698	AC012654 Arabidops
32	71.8	7.3	9424	8	OSU72724	AB010698 Arabidops
33	71.2	7.3	101154	8	ATF1P2	U72724 Oryza sativ
34	70.2	7.2	116205	50	AC006931	U72725 Arabidops
35	69.8	7.1	61712	7	AB017061	AC006931 Arabidops
36	68.6	7.0	13341	8	OLRKLPI	AB017061 Arabidops
37	68.6	7.0	52717	7	AB019227	U72726 Oryza longi
38	68.4	7.0	2278	8	SBRLK1	AB019227 Arabidops
39	68.4	7.0	3573	5	A57133	Y14600 Sorghum b1c
40	68.4	7.0	4000	8	AF053996	A57133 Sequence 4
41	68.4	7.0	5368	8	U42445	AF053996 Lycopersi
42	68.4	7.0	6471	5	A57130	U42445 Lycopersico
43	68.4	7.0	6471	8	U42444	A57130 Sequence 1
44	67.6	6.9	8416	8	OLRKLPI2	U42444 Lycopersico
45	67.6	6.9	111945	8	ATF1C12	U72725 Oryza longi
						AL022224 Arabidops

ALIGNMENTS

[illegible]

QY	541	tttaagccttaatgcaacccgattgacgggtccaatccctagaacatcgaacgtcaatccc	6000
Db	541	TTTACGGCTTATGACAAACCATTTGACGGTCCAAATCCTTAGAGACTACAGGCATCCC	6000
QY	601	aagccttaaaagtlttgacgctccaaagatgtttgttggaaacatcccaacaacg	6600
Db	601	AAAGCTTAAAGTTGTTGACGCTCCAAACCAATGATTTGTGTGAAACAATCCCAACAAACGG	6600
QY	661	acccttggtaacatctcctttacagaacctttggaaacaacccgagattggggggaccgga	7200
Db	661	ACCCCTTGCTCAATCTCTTACAGAACTTTGAGAAACAACCCGAATTTGGAGGACCGGA	7200
QY	721	attacacgggtcttggaaagctacgacatactgacacttgaacaaactgycgaacactgaa	7800
Db	721	ATTACGTGGGTCTTGGACAGCTAGACACTTAACGTGCACCTGAAACAACCTGGCAAAACCTGAA	7800
QY	781	aatgaagaatttgggggtgtgacctgttaagaacacttccaccaattatcaaatccaatc	8400
Db	781	AAAGAAGAATTGGGGGGGACCTCTTGAAGAACCTTCACCACTTTATCAAAATATCACATC	8400
QY	841	tattatgtaataagacatactatgtagttaaaaaaacaataaataagaagatcgaaatcgta	9000
Db	841	TATTATGTAATAAGATATATATATGTAGTAAAAAACAAAAAATAAGAAAGATCGAATCGGTA	9000
QY	901	atatcalctgtgtctcaatttggaaacacttcgagtcgtatgtataaatttcctaattgcgatt	9600
Db	901	ATAATCATCTGGCTCAAAATGTAGAACTTCGAGGCTGTATGTAAATTTCTTAAATGCCGATT	9600
QY	961	tttggccttaattactcaact 981	
Db	961	TTTGCCCTAAATTACTCACACT 981	

[illegible]

```

Db 51 TATTTCGTTTCCTTCCTTTCCTTCAGAAATTTCCCTTACTCTCAAAATTCCTTTTCG 110
OY 73 atttcctctcttaaacctccgaaagctcacatgctgctcgaacatctcgttggagct 132
Db 111 ATTTCCTCTCTTAACCTCCGAAAGCTCACATGCGCTCGAAACTATCGTGGAGCT 170
OY 133 ctctgcaactctgttaaaccttaactgcttgaattcactgctgctgaagaacatccga 192
Db 171 CTTCGCACTTCGTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAAC 230
OY 193 aagagatgctctcctcagctctcgcgcgagttgacagatccacacatgctcctcagag 252
Db 231 AGGAATGCTCTCTACGCTCTTCGCGGAGTTTGACAGATCCACATGCTCCCGAG 290
OY 253 ctggagatccaaactctgttaactcctgtacactggttccatgctcgttaacaaagaca 312
Db 291 CTGGGATCCAACTCTGTAACTTCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 350
OY 313 cggagctcactgctggtggttgggaaattcaaacctcctcgaatctctgscctgagct 372
Db 351 CCGGCTCACTCTGTGGATTGGGAATTCAACTCTCTGACATCTTGCGCTGAGCT 410
OY 373 tgggaagcttgacacttaccagatctagagctctacaaacacacacacacaggaactat 432
Db 411 TGGGAAGCTTGAACATTAACATCTAGAGCTTACAAAACAAACATCCAGGAATCTAT 470
OY 433 accctccgaacttggaatctgaagaatctatagcttgaattgttaacaaacacatc 492
Db 471 ACCCTTCGGAATTTGAATCTGAAGATCTCACTGAGCTGATGTAACAAACAACTCT 530
OY 493 taacaggaatgctccaaactctcttgggaaattgaagctcctggtctcttcttgaagctaa 552
Db 531 TACAGGAGATGTTCCACACTTCTTGGAATTAAGATCTGCTGCTGCTGCTGCTGCTGCT 590
OY 553 tgacaaacgattgacccgctccaaactcctcagagcactcgcgcaatcccaagcc-ttaa 611
Db 591 TGAACAACCATTTGACCGGCTCAAT-CTTAGACACTCAAGGCAATCCCAACCTTTAAG 649
OY 612 ttgttgaagctcgaacatgattgtgtgagacatcccaacaaacgaaccccttctc 671
Db 650 TTGTTGACGCTCAAGCAATGATTGTGTGGACAAATCCCAACAAAGGACCTTCTCTC 709
OY 672 acatccctcttaacaaacttgagaacaaacgagatttggaggagcaggaattactcgtc 731
Db 710 ACATTCCTTTACAGAACTTTGAGAACACCCGAGATTGAGAGACCGGAATTAATCTCG 769
OY 732 ttgcaagctcgaacactcagctcagctgaacaaacttgcaaaccttgaaatgaagaatt 791
Db 770 TTGCAAGCTACGACACTGACACTGAAACACTGCAAAACCTGAAATGAAAGATT 829
OY 792 ggggggtgacacttgaagaacacttcaacacttcaacatatacatctatgtat 851
Db 830 GGGGGGTGACCTTGTAGAACACTTCAACCTTATCAAAATTCATCTTATGTAAT 889
OY 852 aagtatatatgtagttaaaaaaagaaatgaagaatgtaattgtaatatctcgt 911
Db 890 AAGTATATATATGTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 949
OY 912 tctcaattgagaactcgaagctctgtatgttaaatctcaaatgcaatttcgcctaa 971
Db 950 TCTCAATGAGAACTTCGAGGCTCTGTATGTAATAATTCATAATGAGATTTCCGCTACTG 1009
OY 972 tact 975
Db 1010 TAAT 1013

```

```

RESULT 3
LOCUS A67825 1063 bp DNA
DEFINITION Sequence 30 from Patent WO9743427.
ACCESSION A67825
VERSION A67825.1 GI:4756647

```

```

KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
AUTHORS 1 (bases 1 to 1063)
TITLE De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
JOURNAL PRODUCTION OF APOMICRIT SEED
PATENT: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
location/Qualifiers
source 1..1063
/organism="unidentified"
/db_xref="taxon:32644"
106..762
CDS
/note="unamed protein product"
/codon_start=1
/protein_id="CAB4252.1"
/db_xref="GI:4756648"
/translation="MASRNYRMELEFASLILALILHLVYANSSEGDALVALRSLTDP
DHYLSMDPTLVNRCWPHVTCNODRVRVRLDLSNLSNGHAPLGLKLEHLYELY
KNNIQGIPELSENLKMLISLSDLYNNNNLGLYPTSLGKLSVPLRLNDRNLGPTR
ALTRIPSLKVVDSNDLGGTITPTNPFAPHPLPQNFENPRLEGPGLGLASTYDICT
"
BASE COUNT 313 a 242 c 206 g 302 t
ORIGIN
Query Match 85.8%; Score 841.8; DB 5; Length 1063;
Best Local Similarity 96.1%; Pred. No. 2,8e-197;
Matches 885; Conservative 0; Mismatches 32; Indels 4; Gaps 2;
OY 57 cccaattcctttcgaattccctctcttaaacctcgaagctcagatgctcga 116
Db 59 CTCGAATTCATTAATGATTACTCTCTTGACCTCCGATAGCTCACATGCGCTCGAA 118
OY 117 actatcggtggagctcttcgagcttcgttaaaccttaactgaacttgaattcacttg 176
Db 119 ACTATCGGTGGAGCTCTTCGACACTTCGTTAACTTAACCTTAACCTTAACCTTAAC 178
OY 177 tgaagcaaacctcgaagagatgctcctcagctcttcgcgagttgacagatccag 236
Db 179 TCGAAGCAAACTCCGAAAGAGATGCTCTTACGCTCTTCGCGGAGTTTAACGATCCG 238
OY 237 accatgctcccgagagctcggatccaaactctgttaactccttgactcgttccatgta 296
Db 239 ACCATGTTCTCCAGAGCTGGGATCCAACTCTTATCTTACCTTGAACCTGTTCCAGTCA 298
OY 297 cctgttaacaaagacacgcgctcactcgtgtgatttgggaaattcaaacctctcggac 356
Db 299 CCTGTACCAACAAACCGCGTCACTGCTGTGGATTGGGAATTCAACTCTCTGGAC 358
OY 357 atcttgccctgagcttggagacttgacaattacagtaactagagctctacaanaaca 416
Db 359 ACTTGCGGCTGAGCTTGGGAAGCTTGAACATTACAGTATCTAGACCTCTACAAAAACA 418
OY 417 acatcaaggaactatcactccgaacttggaaatctgaagaatctcataagcttggatc 476
Db 419 ACATCCAAAGAACTATACCTTCCGAACCTTGAAGAACTGAAACCTCATAGCTTGATC 478
OY 477 tgtacaacacacatcttcaaggaatgtccactctcttgggaaattgaattcctcgtg 536
Db 479 TGTACAAACAAGATCTTACAGGAGATGTTCCCATCTCTTGGGAAATTAAGTCTCTG 538
OY 537 tcttttacgcttaatgaacacgattgacgggttccaatccctagagcaactcagggaa 596
Db 539 TCTTTTACGGCTTAATGACAAACGATGACGGGCAATCCCTAGGCACTCACTGCAA 598
OY 597 tcccaagcctlaagcttgaagctcgaagcaatgatttgggaaacatccacaaca 656
Db 599 TCCCAAGCCTTAATGTTGTATGTCTCAACAGATGATTGTGTGGAACAATCCCAACAA 658
OY 657 acggacccttgcatactccttaacgaacttggagaacacccgagattggaaggac 716

```

Db 659 ACGGACCTTTGCTCACAATCTTCTTACAGAACTTGAGAACACCCGAGGTTGGAGGAC 718
Qy 717 CGGAATtactcggtcttgcgaagctacgacactaactgacactggaacacactggaacac 776
Db 719 CGGAATtactcggtcttgcgaagctacgacactaactgacactggaacacactggaacac 778
Qy 777 tgaagaatgaagaatgggggggtgacacctgtgaagaacacttcaacattatcaatata 836
Db 779 tgaagaatgaagaatgggggggtgacacctgtgaagaacacttcaacattatcaatata 838
Qy 837 cctcctatgtgaataagatatatactagtgaataaacaacaaataagaatcgaatc 896
Db 839 CATCTACTGTGAATGAATATATATATGATGATCA--AAAAAATAAGATGATC 896
Qy 897 ggtataatcatctgtctcaatggaactggaactcgaagtc--tgatgtataatcttaaat 954
Db 897 AGTAATATCATCTGCTCAATGAGAACTTGAGGTCTGTATGTAATAATTTCTAAT 956
Qy 955 gggatttcgcttaactact 975
Db 957 GCGACTTTCGCGTACTGTAAT 977

RESULT 4
LOCUS A67821 789 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 26 from Patent WO9743427.
ACCESSION A67821
VERSION A67821.1 GI:4756643
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 789)
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMICRITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)

FEATURES
source Location/Qualifiers
1..789

CDS
/organism="unidentified"
/db_xref="taxon:32644"
/note="unnamed protein product"

BASE COUNT 235 a 191 c 156 g 207 t
ORIGIN
Query Match 75.6%; Score 742; DB 5; Length 789;
Best Local Similarity 96.8%; Pred. No. 1.1e-172;
Matches 757; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 101 cacatggcgctcgaactacgttggagctcttcgacgctcgttaacccttaacctta 160
Db 5 CCACGGCGGTCCGGAAACATCGTGAGAGCTCTTCGACGCTCGTAAATCCATACCTTA 64
Qy 161 gctttgttcaactgtgtcgaagcaaacctcgaagagatgctctcaagctcttcgcgg 220
Db 65 GCTTTGATTCACCTGCGGAGCAAACTCGAGAGAGATGCTTTACGCTCTTCGCCGG 124
Qy 221 agtttgacagatccagacatgtccctccagagctggatcccaactctgttaacctgt 280
Db 125 AGTTTACAAATCCGGACCAATGTTCTCCAGAGCTGGGATCCAACTCTGTAAATCTTGT 184
Qy 281 acctgttcatgtcacctgttaaccagaacacgcgcgtcaactcgttgattgggaat 340

Db 185 ACCGTGTCATGTACCTGTACCAAGAACACCCGTCACCTGCTGATTTGGGAAAT 244
Qy 341 tcaacctctctggaactcttcgcctgagcttgggaagcttgaacatttcaatata 400
Db 245 TCAACCTCTCTGAGATTTTGCGCTGAGCTTGGAAGCTTGAAACATTTACAGATCTA 304
Qy 401 ggaacttcaaaaaaacaatccaaggaactataccttcggaacttgaatctgaagaat 460
Db 305 GAGCTGTACAAAACAACTCCAGGAACTATACCTTCGAACTTGGAATCTGAAGAT 364
Qy 461 ctcatagcttgcgtctgtcaacaacaactcttacaggaatagttccactcttggga 520
Db 365 CTATCAGCTTGATGATGTGACAAACAAATCTTAAGGATAGTTCCACTCTTTGGGA 424
Qy 521 aaattgaatctctgtctctttttagcgtcttaagaacacgattgacccgtccct 580
Db 425 AAATGAACTCTGCTCTTTTACGCTTAATGACAAACGATGACGGGGCCAACTCCCT 484
Qy 581 agagcactcaagcgaatcccaagccttaagtgtgtgaagcttcaagcaatgattgtgt 640
Db 485 AGAGCACTCCTGCAATCCCAAGCCTTAAGTGTGATGTCAAGCAATGATTGTGT 544
Qy 641 ggaacaatcccaacgaagcacccttgcacatctccttaacagaacttgaagaacac 700
Db 545 GGAACAAATCCCAACAAACGACCTTTGCTACATTCCTTACAGACTTGAAGAACAC 604
Qy 701 ccgaagtggagggacgggaataactcgtcttgcgaagctacgaacactgaacctga 760
Db 605 CCGAGGTGGAGGGACCGGGAATTAAGCTGTGCAAGCTACGACAACTAACGCACTGGA 664
Qy 761 aacaactggcaaacctgaaatgaagaattgggggtgacacttgaagaacactcacc 820
Db 665 AAAAATGGCAAAACCTGAAATGAAGATGGGGGTACCTTTGACAACTTCCACC 724
Qy 821 actttatcaaatatcactatataatgaatgaatgaatgaatgaatgaatgaatgaat 880
Db 725 ACTTTATCAAAATATCAGATCTATGTAATGATATATATGATGATGATGATGATGAT 784
Qy 881 aa 882
Db 785 AA 786

RESULT 5
LOCUS A67823 894 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 28 from Patent WO9743427.
ACCESSION A67823
VERSION A67823.1 GI:4756645
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 894)
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMICRITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)

FEATURES
source Location/Qualifiers
1..894

CDS
/organism="unidentified"
/db_xref="taxon:32644"
/note="unnamed protein product"

Qy 101 cacatggcgctcgaactacgttggagctcttcgacgctcgttaacccttaacctta 160
Db 5 CCACGGCGGTCCGGAAACATCGTGAGAGCTCTTCGACGCTCGTAAATCCATACCTTA 64
Qy 161 gctttgttcaactgtgtcgaagcaaacctcgaagagatgctctcaagctcttcgcgg 220
Db 65 GCTTTGATTCACCTGCGGAGCAAACTCGAGAGAGATGCTTTACGCTCTTCGCCGG 124
Qy 221 agtttgacagatccagacatgtccctccagagctggatcccaactctgttaacctgt 280
Db 125 AGTTTACAAATCCGGACCAATGTTCTCCAGAGCTGGGATCCAACTCTGTAAATCTTGT 184
Qy 281 acctgttcatgtcacctgttaaccagaacacgcgcgtcaactcgttgattgggaat 340

RESULT	7
SBUS62279	936 bp mRNA PLN 09-DEC-1996
LOCUS	
DEFINITION	Sorghum bicolor leucine-rich repeat-containing extracellular glycoprotein mRNA, complete cds.
ACCESSION	U62279
VERSION	U62279.1 GI:1710123
KEYWORDS	
SOURCE	sorghum.
ORGANISM	Sorghum bicolor
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumhyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.
AUTHORS	1 (bases 1 to 936)
TITLE	Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.
JOURNAL	Isolation of a cDNA encoding a novel leucine-rich repeat motif from Sorghum bicolor inoculated with fungi
MEDLINE	Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)
REMARK	97124217
REFERENCE	Erratum: [published erratum appears in Mol Plant Microbe Interact 1997 Mar;10(2):3021]
AUTHORS	2 (bases 1 to 936)
TITLE	Hipskind,J.D.
JOURNAL	Direct Submission
FEATURES	Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA
source	Location/Qualifiers
CDS	1..936
	/organism="Sorghum bicolor"
	/db_xref="taxon:4558"
	/clone="SLRR"
	12..735
	/function="protein-protein ligand binding"
	/note="leucine-rich repeat-containing extracellular glycoprotein; contains six N-glycosylation sites [NX(S/W)]"
	/codon_start=1
	/protein_id="AAC49559.1"
	/db_xref="GI:1710124"
	/translation="MAPQFAAGFGLALATASCTEGDILYKORLAWEDNNVYLQSNMSTLANPCTWFHVCNNNNFVRIYDLGNAGISGLPLAIQNIQYELKNGIGNSIPETIGNTNLISDMDNLGTGEPTLGSVSTRYLRKIVNNMTGPIPSFGGLTSLSEKIQENSLSGAIPASIGIKKALQERLMDNMTGVIPSKSPISFGNLTETLMTDRNNDGTRTSSSLRTAIIQDAKRA"
	36..77
	/evidence="not_experimental"
BASE COUNT	247 a 232 c 207 g 250 t
ORIGIN	
Query Match	16.0%; Score 156.6; DB 8; Length 936;
Best Local Similarity	58.5%; Pred. No. 1.6e-28;
Matches 273; Conservative	0; Mismatches 194; Indels 0; Gaps 0;
185	aacctccgaaggagatgctctcagcgtcttcgcgcggagtttgcagatccagaccatgfc 244
81	aacactggaagtgacatctatctgtacaaagcaaaagttggcatgggagaccacaacacg 140
245	ctccagagctggagatccaactctgtttaactccttgtaacctggttccatgtaacctgtaac 304
141	ctgcacagctgggaattcgacgctgtgcgaatcccttcgacacctgttcattctaccttcacac 200
305	caagacaacgcggtcactcgtgtgagtttgggaataattcaacctctctggaatcttcg 364
201	aacatataacttttgtcatccgcgttgattgggcatacgacgcatctccggtccttgc 260
365	cttgagcttgggaagcttgaacattcagatctcagatccatcacaaaacaacatccaa 424
261	ccagactcttgcaagaatttcagaaactccagatccagatccagatccagatccagatccag 320
425	ggaactatactctcgaaacttggaaatctggaagaatctcaacagcttgatctgtacaac 484
321	ggttcatttcacaaacactatgacacactgcacaaactcattcaatctatcagcttgatcttgagac 380

QY	485	aaacatcttaacggatggtccactctcttgsgaaatctgaagtcctgctcttta	544
Db	381	AACCTCTTACCGGGGCAAAATCCAACTACGCTTGTTCTCTCAGCAGCCTGGCATATCTG	440
QY	545	cggcttaatgacacagatctgaccggtccatcccttaagcactcaagcgaatcccaagc	604
Db	441	AGGTTGACCAAAACAACACTGACAGGGCGTATTACCATCATCTGTTTGCAACCTGACTAGC	500
QY	605	cttaagtggttggaagtcctcaagcattgattggtgtaacatcc	651
Db	501	CTTCTGTAATCGAAGCTTCAAGAGATTGTTGAGCGCGCTATTCC	547
RESULT	8		
LOCUS	DCU93048	1755 bp	mRNA
DEFINITION	Daucus carota somatic embryogenesis receptor-like kinase mRNA,		
ACCESSION	complete cds.		
VERSION	U93048		
KEYWORDS	U93048.1 GI:2224910		
SOURCE	carrot.		
ORGANISM	Daucus carota		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphylliphtes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.		
AUTHORS	1 (bases 1 to 1755)		
TITLE	Schmidt,E.D., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.		
JOURNAL	A leucine-rich repeat containing receptor-like kinase marks somatic		
REFERENCE	plant cells competent to form embryos		
AUTHORS	Development 124 (10), 2049-2062 (1997)		
TITLE	97313247		
JOURNAL	2 (bases 1 to 1755)		
REFERENCE	Schmidt,E.D., Toonen,M.A.J. and de Vries,S.C.		
AUTHORS	Schmidt,E.D., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-MAR-1997) Molecular Biology, Agricultural University		
FEATURES	of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands		
SOURCE	Location/Qualifiers		
	1..1755		
	/organism="Daucus carota"		
	/db_xref="taxon:4039"		
	34..1695		
	/note="SERK"		
	/codon_start=1		
	/product="somatic embryogenesis receptor-like kinase"		
	/protein_id="AAB61708.1"		
	/db_xref="GI:2224911"		
	/translation="MNRSINILNMOFTDAYLDKYGVLMTELEKSNNSISGPISPDLDL		
	NLTNTVSIDVIMNSFGPIPTDGLKTRLEFLFNNSISGPIPSLNTITVILVLDL		
	PNNRSISGVPDNGSPSLFTPIPSANNLTGCPVTGRCGSPSPSPPIIPSPVBO		
	PQONPETHINGVAAAGAAALFAAPAMAFMMRRRRPRHFVDVRAEDPEYHAGQLK		
	RSLSELDVADTFTSLIRGGEFGKVKGLAGLSVAVKRLKEETPEGLQFOTE		
	EKLISAVVHNLRLFGMCPMTPTERLVLVPPMANGSVASGLRRGSPSEPLMPTKTR		
	AAGSARGISYLDHDDCPILIRHDVKANILILDEPAAVVDGLRLANDYKDTHTA		
	VGGTGLIAPETLSTGKSEKTEIVFGIMLLILITGSAFPLDRLANDDDVMLDWM		
	KSLKEKLEMLPDDLENNITDEYVQLIIVALLITQSPSPERPMSVYRMLEGDC		
	LAEKDEMKVEYVHODELAPHTFSTMIIDSLNLAHELISGR"		
BASE COUNT	506 a	347 c	407 g 495 t
ORIGIN			
Query Match	10.4%	Score 101.8;	DB 8; Length 1755;
Best Local Similarity	55.1%	Pred. No. 5.1e-15;	
Matches 199; Conservative	0;	Mismatches 162;	Indels 0; Gaps 0;
QY	361	tgcgcctgaagcttgggaagcttgaacattcaagatcctagactctacaaacaacat	420
Db	78	TGATGCTTACCTTGACAAATATATGGGGTCTTATGACATTGAGCTTACAGCATAAATACAT	137
QY	421	ccaaggaactatcctctcgaacttggaaatctgaagaatctatcgcgttgatctgta	480
Db	138	AAGTGACCAATTCCTAGTATCTTGGGAAATCTGCAAAATTTGGTGTGAGCTTGACCTATA	197

QY	481	caacaaacaaactcttaagggatagatgccactctcttgggaaaaatcgagctctcgagctc	540
Db	198	CATGAATACGCTCTCTGGACCTATACCGGACACATTAGAAAGCTTACAAAGCTTAAGATT	257
QY	541	tttagcgcttaatlgacaacgcatctgacccggtlccaatcccttagagcactacagcgcaatcc	600
Db	258	CTTGCGCTTCACACAAACAGCGCTCTCGGCCAATTTCCAAATGTACGTACATTAATTATAC	317
QY	601	aagcgcttaagttgttgacgcgtcctgaagcaatgattgyltggaacaatcccaacaacg	660
Db	318	AACCTTCAAGTCCTGGATTATACAAACATCGGCTATCAGGACCAATGACCGGATTAATGG	377
QY	661	acccttgcctcaatctcctttagaagaaacttggaaacaacccgagatcttggagggaccgga	720
Db	378	CTCATTTCTCTTGTTTACACCATACAGTTTGGCAATTAATTGAAATTATATGAGCCCGT	437
QY	721	a 721	
Db	438	A 438	

RESULT	9			
LOCUS	A67797			
DEFINITION	Sequence 2 from Patent WO9743427.			
ACCESSION	A67797			
VERSION	A67797.1			
KEYWORDS	GI:4756623			
SOURCE	Carrot.			
ORGANISM	Daucus carota			
REFERENCE	Daucus carota			
AUTHORS	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
TITLE	eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.			
JOURNAL	1 (bases 1 to 1815)			
	De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.			
	PRODUCTION OF APOMICRITIC SEED			
	Patent: WO 9743427-A 20-NOV-1997;			
	CIBA GEIGY AG (CH)			
FEATURES	Location/Qualifiers			
source	1..1815			

```

/organism="Daucus carota"
/db_xref="taxon:4039"
94. .1755
CDS
/note="unnamed protein product"
/peptide_start=1
/peptide_end=1
/peptide_id="CAB42241.1"
/db_xref="GI:4756624"
/translation="MNRMSILNMYQFTADLYDKYGLMTLETLSNNISGPIPSIDG
NITNYSIDLYNNSGSPIDPTGLKRLRLPGLTNNNSISGIPSLNITNITLVIDL
NNRSLGSPYPNQSPSLPTSPANNINLTGPGTGPCGSPSPSPPPPIIPSTPQ
RQNPPTGATNGVAAAGAILLAPAPAFMMARRRRRPHHFDVQAEEDPEHISGLQ
RPSRLQVADYDIPSTILRGQFGKRYKMGAGSIVAKRLKESTPEPGLQGTETE
KHSIAVERNLRLRGCFMPTERTLVLPYMGASVAVAKRLKESTPEPGLQGTETE
ALGSGAGSLYHDCDPRILHRDVKANILIDEEFVAGVGLRLANDYDVTHTTA
VGTGGLYAPETLSGKSEKTDVFGYINLLDTTGQAFPLARLANDYDVTHTTA
KSLTLERKLEMLVDPLENLENNITDTEVEQLIQVALLCTQSPPERKMSSEVYMLDEGQ
LAEKDEWQKVEYIHQDELAPLHAPSLMIDSTYDNIHAPSLSPGR"
531 a 354 c 415 g 515 t
BASE COUNT
ORIGIN

```

	Query Match	Best Local Similarity	Matches	Conservative	Score	DB	Length	Mismatches	Indels	Gaps
QY	361	tcgcgccttgcgagcttgcgagcttgaacattgaacatgaagcttcacaaagaacat	420		10.4%	101.8	480	5	1815	
DB	138	TGATGCTTACCTTACCAATATGAGGCTTTTATGACATTGAGCTTTGACCAATATAC	197		15.1%	5.1-15				
QY	421	ccaaggaactataacttcgcgaacttgcgaacttcgaagaacttcacatgcgcttgatctgta	480							
DB	198	AAGTGACCAATTCCTTCATGATCTTTGGGATCTGACCAATTTGGATGGAGCTTGACCTTAA	257							

QY	481	caacaacaatcttaaggagatagttcccaactctttgggaaatagagtcctcggcct	540
Db	258	CATGATAGCTTCTCTGGACCTATACGGACACTTTAGGAAGCTTTACAGGCTTAGATT	317
QY	541	tttagcgcttaatgacaacagatgaccggctccaatcccttagagactcaacgacaatcc	600
Db	318	CTTGCGTCTCAACACAAAGAGCTCTCGTGCAATTCGAATGTCACTGACCTAATATTAC	377
QY	601	aagccttaaaagtctgtgacgcctcaagcaatgattgtgtggaacaatcccaacaacgg	660
Db	378	AACCTCTCAAGTCTCGGATTTATCAACAATCGGCTATCAGGACCGATCCGGATATGG	437
QY	661	acccttgctcaatcctcctttacagaactcttgaaacaacccgagatcttgaggagacgga	720
Db	438	CTCATTTTCTTGTATTACACCATACGTTTGGCAATTAATTAATTAATGAGCCCGT	497
QY	721	a 721	
Db	498	A 498	

RESULT	10				
LELRGENE	LOCUS	LELRGENE	4604 bp	DNA	PLNT
DEFINITION		L.esculentum LRP gene.			
ACCESSION		X85269			
VERSION		X85269.1	GI:1619299		
KEYWORDS		LRP gene; LRR protein.			
SOURCE		tomato.			
ORGANISM		Lycopersicon esculentum			
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; Solananceae; Solanales; Solanaceae; Solanum; Potatoes; Lycopersicon.			
REFERENCE		1 (bases 1 to 4604)			
AUTHORS		Vera, P.			
TITLE		Direct Submssion			
JOURNAL		Submitted (18-JAN-1996) P. Vera, Universidad Politcnica de Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino de Vera 14, E- 46022 Valencia, SPAIN			
REFERENCE		2 (bases 1 to 4604)			
AUTHORS		Tornero, P., Mayda, E., Gomez, M.D., Canas, L., Conejero, V. and Vera, P			
TITLE		Characterization of LRP, a leucine-rich repeat (LRR) protein from tomato plants that is processed during pathogenesis			
JOURNAL		Plant J. 10 (2), 315-330 (1996)			
MEDLINE		96367673			

FEATURES	source	location/Qualifiers
exon		1..4604
gene		/organism="Lycopersicon esculentum"
		/cultivar="VFN8"
		/db_xref="taxon:4081"
		/tissue_type="leaf"
		<1..345
		/number=1
		113..4355
		/gene="lrp"
CDS		join(113..345,2381..2452,2565..2708,2928..2999,3081..3152,4283..4355)
		/gene="lrp"
		/codon_start=1
		/product="LR protein"
		/protein_id="CA64565.1"
		/db_xref="GI:1619300"
		/db_xref="SPTREMBL:Q96477"
		/translation="MEAVFKTQSLFLKIMGLAAVAVAKGNSGDALFYALRSLS
		SDGNAVSDWPNLVNPTGTFVTCNGNQVTRVDIGSKLSGLHLYPELGLEHLYQL
		ELKKNNGOIGIPKEINGLKSLISLDLYNNISGRTPTSLGMLKNIVFLRLNDKNTLP
		IIPRLNLSISLAKVVDVSNNDLCGTIPTSQPFPHILNNFENRPLEGEELLGLATYDT
		NC5"
intron		346..2380
		/gene="lrp"
		/number=1

exon	2381..2452	/gene="LRP"
	/number=2	
Intron	2453..2564	/gene="LRP"
	/number=2	
exon	2565..2708	/gene="LRP"
	/number=3	
Intron	2709..2927	/gene="LRP"
	/number=3	
exon	2928..2999	/gene="LRP"
	/number=4	
Intron	3000..3080	/gene="LRP"
	/number=4	
exon	3081..3152	/gene="LRP"
	/number=5	
Intron	3153..4282	/gene="LRP"
	/number=5	
exon	4283..>4604	/number=6
	/number=6	
BASE COUNT	1311 a 758 c 888 g 1647 t	
ORIGIN		

Query Match	9.8%	Score 96.6	DB 7	Length 4604
Best Local Similarity	69.1%	Pred. No. 9,9e-14		
Matches 137	Conservative 0	Mismatches 59	Indels 0	Gaps 0
Qy	140	gcttcgttaacccttaaccttagcttgatcaccctgctcgagaacaaacctccgaagagat	199	
Db	158	GGTTTGTGGCAATTGTCCTGTGACGTGAGCTGTAAAGGGGGAATTGAAAGGGGAT	217	
Qy	200	gctctcaagcctcttcgcgcgagattgtagacagatccacagacccatgtccctccagagctggat	259	
Db	218	GCTTTGTAGCCCTTGTGGCCGGAGCTTATGTGACCCGGGTAAAGTGTAAACAGAGCTGGGAT	277	
Qy	260	ccaactctgttaatccttgtaacctggtacatgtccatgtccacttaaccaagacaaccgctc	319	
Db	278	CCAAATGTTTGTATACCTGTGTACCTGTGCTTCAATGTCACCTGCACAGGAGATTAACAAGTT	357	
Qy	320	actcgtgtgga	330	
Db	338	ACTCGTGTGTA	348	
RESULT 11				
LOCUS	ATAC009991/c			
DEFINITION	Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence.			
ACCESSION	ATAC009991			
VERSION	AC009991.3			
KEYWORDS	GI:601677			
SOURCE	HTG.			
ORGANISM	thale cress			
	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;			
	Arabidopsis.			
REFERENCE	1 (bases 1 to 101284)			
AUTHORS	Lin,X., Kaul,S., Town,C.D., Benito,M., Greasy,T.H., Haas,B.,			
	Roemmig,C.M., Koo,H., Fujii,C.Y., Utehrack,T.R., Barnstead,M.E.,			
	Bowman,C.L., White,O., Mierman,W.C. and Fraser,C.M.			
TITLE	Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 101284)			
AUTHORS	Lin,X. and Kaul,S.			

TITLE	COMMENT
Direct Submission	
Submitted (09-SEP-1999) The Institute for Genomic Research, 9712	
Medical Center Dr, Rockville, MD 20850, USA, xlinetligr.org	
3 (bases 1 to 101284)	
Lin.X.	
Direct Submission	
Submitted (08-OCT-1999) The Institute for Genomic Research, 9712	
Medical Center Dr, Rockville, MD 20850, USA	
On Oct 8, 1999 this sequence version replaced g1:5902413.	
Address all correspondence to:	
Xiaoying Lin	
The Institute for Genomic Research	
9712 Medical Center Dr.	
Rockville, MD 20850, USA	
e-mail: xlinetligr.org	
BAC clone Fg98 is from Arabidopsis chromosome III and is near the	
molecular marker g4547.	
The orientation of the sequence is from SP6 to T7 end of the BAC	
clone.	
Genes were identified by a combination of three methods: Gene	
Prediction programs including GRL (available by anonymous ftp	
from arthur.emb.ornl.gov), GeneFinder (Phil Green, University of	
Washington), Genscan (Chris Burge,	
http://genomic.stanford.edu/~chris/GENSscan.html), and NetPlantGene	
(http://www.cds.dtu.dk/netcpene/cbsnecpene.html), searches of the	
complete sequence against a peptide database and the Arabidopsis	
EST database at TIGR (http://www.tigr.org/tdb/at.html).	
Associated genes are named to indicate the level of evidence for	
their annotation. Genes with similarity to other proteins are named	
after the database hits. Genes without significant peptide	
similarity but with EST similarity are named as 'unknown' proteins.	
Genes without protein or EST similarity, that are predicted by more	
than two gene prediction programs over most of their length are	
annotated as 'hypothetical' proteins. Genes encoding tRNAs are	
predicted by tRNAscan-SE (Sean Eddy,	
http://genome.msu.edu/eddy/tRNAscan-SE/). Simple repeats are	
identified by repeatmasker (Arjan Smit,	
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of	
genomic sequence that are not annotated as genes but have predicted	
exons by GRL are annotated as misc features.	
Location/Qualifiers	
1..101284	
/organism="Arabidopsis thaliana"	
/cultivar="Columbia"	
/db_xref="taxon:3702"	
/chromosome="III"	
/map="g4547"	
/clone="Fg98"	
77..108	
/note="exon predicted by xgrall, quality excellent"	
124..259	
/note="exon predicted by xgrall, quality excellent"	
complement(347..408)	
/note="exon predicted by xgrall, quality	
excellent_shadowexon"	
355..594	
/note="exon predicted by xgrall, quality excellent"	
672..752	
/note="exon predicted by xgrall, quality excellent"	
854..973	
/note="exon predicted by xgrall, quality excellent"	
complement(1182..1357)	
/note="exon predicted by xgrall, quality	
marginal_shadowexon"	
complement(1356..1664)	
/note="exon predicted by xgrall, quality	
excellent_shadowexon"	
1546..1666	
/note="exon predicted by xgrall, quality good"	
1747..1838	
/note="exon predicted by xgrall, quality excellent"	
complement(2090..2250)	
/note="exon predicted by xgrall, quality good_shadowexon"	

misc_feature	2114..2246	/note="exon predicted by xgfall, quality excellent"
repeat_region	complement(2975..3024) /rpt_family="(CAA)n"	
mRNA	join(3070..3325,3565..3642,4094..4234,4338..4414, 4506..4655,4800..5013) /gene="P9F8.1" 3070..>5013	
gene	/gene="P9F8.1" /note="similar to unknown protein GB:AAC27167 [Arabidopsis thaliana]"	
CDS	join(3215..3325,3565..3642,4094..4234,4338..4414, 4506..4655,4800..5013) /gene="P9F8.1" /note="unknown protein" /codon_start=1 /protein_id="AAFO1505.1" /db_xref="GI:6016678" /translation="MVGAPQIVLFGSSIVQSGFGHGWALISEYARKADIIILNG KYSMSRLAEVDQVFPKDAVQPSLIVYFEGGDSMAHPSGLGPHVLEVDNM KRIALHISLSPFRRIIFLSPDPEAKRONSPYLSEYLRNDCIKTYSDACVELC QELGLEVDLESTFOKADDKTKVCTDCHLSAQSKITVAGEIILRVKEDMNPSLHM KSMPEFADSPDSYDVLVSADGQTVNSSEWTFMEQWD"	
repeat_region	complement(5513..5597) /rpt_family="(GA)n"	
misc_feature	complement(5527..5579) /note="exon predicted by xgfall, quality excellent shadowexon"	
mRNA	join(5588..5795,5897..5929,6013..6235,7264..7387, 7467..7932) /gene="P9F8.2"	
gene	5589..7932 /gene="P9F8.2" /note="similar to nucleic acid binding protein Alfin-1 GB:AAA20093 [Medicago sativa]"	
CDS	join(5696..5795,5897..5929,6013..6235,7264..7387, 7467..7727) /gene="P9F8.2" /codon_start=1 /product="putative nucleic acid binding protein" /protein_id="AAFO1506.1" /db_xref="GI:6016679"	
repeat_region	/translation="MAAAVSSNPRTVEIRFDYSARRAALLRLTKDVDFYSQCP EKEELNGLPHEWSEVNLPAEVEPELPLGINFARDQKRDWISLVAHSDGW LISVSYFARLNRNEKRLFLINDPLTEPVYGRAMDNRSPDSGSKNGTK RLSDGQTKSTPKLMESEYEEEDDEHDGDLCSGGHYNEEFWICDVCERWYHG KCKIRPAKESIKYKCCPCCAKKGRC"	
repeat_region	complement(7717..7760) /rpt_family="(CAT)n"	
repeat_region	complement(7810..7847) /rpt_family="(GAAA)n"	
mRNA	8515..8572 /rpt_family="(TAA)n" complement(join(9426..9674,10198..10528,11098..11345, 11491..>11865)) /gene="P9F8.3"	
gene	complement(<9426..>11865) /gene="P9F8.3" /note="similar to leucanthocyandin dioxigenase GB:BA420143 [Perilla frutescens]"	
CDS	complement(join(9426..9674,10198..10528,11098..11345, 11491..11865)) /gene="P9F8.3" /codon_start=1 /product="putative leucanthocyandin dioxigenase" /protein_id="AAFO1507.1" /db_xref="GI:6016680"	
repeat_region	/translation="NNNLDIKIESKTCANDOEYKIDNMMSDQKNKIEIKKSG LGKRWPIYRQVSLASNLSTSLPDYIKPSPORQTTIDHPVADINIIIDLOS LFGSNDKGRKRISEACREMGROYINHGKRPILMDAARETMSFNLVPEAKYNS PRTEGEGDKRISEAKREMGROYINHGKRPILMDAARETMSFNLVPEAKYNS LRTVEGDSKRISEAKREMGROYINHGKRPILMDAARETMSFNLVPEAKYNS GKIGRLMTLLSSNGIRAEQDLAEAGGSDVACLRNVPFCQPELALISPSHSD VGTIGRLMTLLSSNGIRAEQDLAEAGGSDVACLRNVPFCQPELALISPSHSDV	

	repeat_region	NSEKERSIAEFVYMPKSDIPDQMGLQWLTSTMBPLYPWTFFQYRLFIIRTGPGRGKSGF VESHSRSP"
	misc_feature	11962..11996 /rpl_family="(GA)n"
	misc_feature	12824..12836 /note="exon predicted by xgrall, quality marginal_shadowexon"
	misc_feature	complement(13367..13468) /note="exon predicted by xgrall, quality marginal"
mRNA	misc_feature	complement(13742..13788) /note="exon predicted by xgrall, quality marginal"
	gene	complement(join(13822..14179,14261..14398,14502..14582,14676..14861,14941..15033,15124..15190,15272..15361,15662..16155))
		/gene="FgP8.4"
		/codon_start=1
		/product="Omega-3 fatty acid desaturase, chloroplast precursor"
CDS		/protein_id="AAP01508.1"
		/db_xref="GI:6016681"
		/translation="MANIVLSEGCIRLPITYTPPSNFTSNNKKRPFLSSSKRTT SSSLPEGLNSRDGEFRNALNVSTPLTPPIFEESPLEDNKQFDGAPPPNLADII AAIPIHGWKNPKMSLSIYVDVAIVPALAAGAYLNMYVPLYLWLAGTFEMALFY LCHDCGSFWSNDPILNSVGVHLHLSIILPYIGWMRIISRTHQNHGHVENDESHMPP SKRIYNTLDKPIREFRFTLPVLMAIYPFLMARSPPKSGHYHPDSIFLPERRDVI

Query Match 8.5%; Score 83.2; DB 8; Length 101284;

Best Local Similarity 54.6%; Pred.No.2e-10;

Matches 166; Conservative 0; Mismatches 138; Indels 0; Gaps 0

Oy	348	tctctggacattctgcgcctagctgagcttgggaacttggaaccattcaagtatctagaagcttc	407
Db	44936	TCTTTGGTCAGATTCCATCTCGATTGGAGAACCTTGCICGTCACTCACTCATCTCTACTTT	44877
Oy	408	acaaaacacaatccaaaggaactataccttcggaaacttggaaacttgaagaattcatca	467
Db	44876	CTTACAAACAATTCGTGGTGGAAATCCCATCTCTTTTGGCAATCTAAACCGCTGATCG	44817
Oy	468	gcttgagatcgtgtaacaacaatatcttaacagggaatagttcccactctcttggaaaatat	527
Db	44816	TCTTACAGGTTGATTCCAATAAAGCTAGAGTAAGTATGCCATCTCAGCTACTGAAATTTGA	44757
Oy	528	agtcctcgatctttttaacggtctaatagatgaacaacggatlgaccggtccaatccctagaagcac	587
Db	44756	CGAGGTTGTCTGCTTTATCTTATCTTCCCAACAACAGTACACAGGCGAGTATTCCTAATACA	44697
Oy	588	tcaaggccaatccccaaagccttaaagttggttgaaagctctcaagcaatgatatttgttgtaacaa	647
Db	44696	TCAAGTTTCTATCAAACCTGATGAGCTTTGGAAGACAGTAGTACACAGCCTTTCAGTGAATCTC	44637
Oy	648	tccc ccc 651	
Db	44636	TCCC 44633	

RESULT 12

AC021198 AC021198 81513 bp DNA PLN 09-FEB-2000

DEFINITION Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete

ACCESSION AC021198

VERSION AC021198.2 GI:6957696

KEYWORDS Htg.

SOURCE thale cress.

ORGANISM	Arabidopsis thaliana Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllales; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.					
REFERENCE	1 (bases 1 to 81513)					
AUTHORS	Liu S.X., Sakano,H., Yu,G., Lee,J.M., Lenz,C., Pham,P., Toriumi,M., Chin,C., Chlou,J., Choi,E., Chung,M., Gonzalez,A., Homg,B., Llu,A., Vaysseier,M., Altali,R., Brooks,S., Buehler,E., Chao,Q., Com,L., Conway A.B., Hansen,N.F., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J., Shin,P., Southwick,A., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.					
TITLE	Arabidopsis thaliana chromosome I BAC F14D7 sequence					
JOURNAL	unpublished					
REFERENCE	2 (bases 1 to 81513)					
AUTHORS	Theologis,A.					
TITLE	Direct Submission					
JOURNAL	Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA					
REFERENCE	3 (bases 1 to 81513)					
AUTHORS	Theologis,A.					
TITLE	Direct Submission					
JOURNAL	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA					
COMMENT	On Feb 11, 2000 this sequence version replaced gi:6693723. The sequence is of BAC F14D7 from Arabidopsis thaliana chromosome 1. The sequence is does not represent the sequence of the entire insert of this clone. It is shorter by 6954 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide small overlaps (200 bp) between overlapping submitted clones. The 5' end of this sequence overlaps by 200 bp to the 3' end of the sequence of the clone F15O4.					
FEATURES	Location/Qualifiers					
SOURCE	1..81513 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="1" /clone="F14D7"					
BASE COUNT	27779 a 13814 c 13668 g 26052 t					
ORIGIN						
Query Match	8.3%; Score 81.4; DB 50; Length 81513;					
Best Local Similarity	54.1%; Pred. No.5,6e-10;					
Matches 166; Conservative	0; Mismatches 141; Indels 0; Gaps 0;					
OY	345	acctctctygaacatcttgcgtcgttagcttgggaaagcttgaacatttacgltatctagaac	404			
DB	2466	ACTTAACTGTGGCATTTCCCGGAACCTGAGCAACAATAATCCAAGATCGATTNGAOT	2525			
OY	405	tctacaacaacaacatccaaggacataacccttcgaacttggaactcgaaatatca 464				
DB	2526	TGAGTAACAACAACCTTACC GGTTCAATTCCTTCTCTTAGAANAATCCAAGAATTGGA	2585			
OY	465	tcaagcttgatctgtacaacaacaacatctcacagagatgtccaccattcttgggaaat	524			
DB	2586	CYATTCTTATTCTTTTGCGAGAACTATTTAACGTGTCATTCGCCCGGAACCTGAGCAACA	2645			
OY	525	tgaagctcttgtctctttacagcgctaagacaacacgatgacacggttcacatccctagag	584			
DB	2646	TGGAAATCGATGTGATTTCAGTTACGATTAAATRAACAACAACCTTACC GGTTCAATTCCTT	2705			
OY	585	cactcagcgcaatccccaaagccttaaaagtgtgtacgltccaagcaatgatttgtggaa	644			
DB	2706	CCYTTCGAATATCAGAACTTGACGTATCTTATCTTTACCTGAATTATTTAACGTGTC	2765			
OY	645	caatccc 651				
DB	2766	TCATTGC 2772				

```

RESULT 13
LOCUS ATAC011765
DEFINITION Arabidopsis thaliana chromosome I BAC F1M20 genomic sequence,
complete sequence.
ACCESSION AC011765
VERSION AC011765.4
KEYWORDS GI:6539234
SOURCE
ORGANISM
Arabis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidops.
REFERENCE
1 (bases 1 to 134402)
Lin,X., Kaul,S., Town,C.D., Beito,M., Crosby,T.H., Haas,B.,
Ronnberg,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
Arabis thaliana chromosome I BAC F1M20 genomic sequence
Unpublished
2 (bases 1 to 134402)
Lin,X. and Kaul,S.
Direct Submission
Submitted (14-OCT-1999) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 134402)
Lin,X.
Direct Submission
Submitted (08-DEC-1999) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Dec 8, 1999 this sequence version replaced gi:6102640.
Address all correspondence to:
Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr.
Rockville, MD 20850, USA
e-mail: xlin@tigr.org
BAC clone F1M20 is from Arabidopsis chromosome I and is near the
molecular marker ml425.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL (available by anonymous ftp
from arthur.spm.ornl.gov), GeneFinder (Phil Green, University of
Washington), Genscan (Chris Burge,
http://www.genomic.stanford.edu/~chris/GENSCANW.html), and NetplantGene
(http://www.obs.dtu.dk/netgene/obsnetgene.html), searches of the
complete sequence against a peptide database and the Arabidopsis
EST database at TIGR (http://www.tigr.org/cdb/at/est.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown'. Proteins
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as 'hypothetical'. Proteins. Genes encoding tRNAs are
predicted by tRNAscan-SF (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SF/). Simple repeats are
identified by RepeatMasker (Arian Smil,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
genomic sequence that are not annotated as genes but have predicted
exons by GRAIL are annotated as misc features.
location/Qualifiers
1..134402
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="I"
/map="ml425"
/clone="F1M20"
138..269
misc_feature
138..269

```


BASE COUNT	FEATURES	COMMENT	TITLE	JOURNAL	REFERENCE	ORGANISM	KEYWORDS	ACCESSION	LOCUS	RESULT
35310 a 20450 c 20573 g 36044 t 152 others	source	On Jan 12, 2010 this sequence version replaced gi:6437511.	Direct Submission	Submitted (16-NOV-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA	Unpublished	Arabidopsis thaliana	HTG; HTGS_PHASE1.	AC015446	AC015446/c	12-JAN-2000
		* NOTE: This is a 'working draft' sequence. It currently								
		* consists of 4 contigs. The true order of the pieces								
		* is not known and their order in this sequence record is								
		* arbitrary. Gaps between the contigs are represented as								
		* runs of N, but the exact sizes of the gaps are unknown.								
		* This record will be updated with the finished sequence								
		* as soon as it is available and the accession number will								
		* be preserved.								
		1 3149: contig of 3149 bp in length								
		* 3150 3199: gap of unknown length								
		* 3200 12898: contig of 9699 bp in length								
		* 12899 12948: gap of unknown length								
		* 50755 50757: contig of 37808 bp in length								
		* 50806: gap of unknown length								
		* 50807 112529: contig of 61723 bp in length.								
		Location/Qualifiers								
		1. 112529								
		/organism="Arabidopsis thaliana"								
		/db_xref="taxon:3702"								
		/chromosome="1"								
		/clone="F12612"								

ORIGIN

Query March 164: 8.2%; Score 80; DB 44; Length 112529; Predicted Local Similarity 53.9%; Pred. No. 1,36-09; Matches 164; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY	327	tgagattgggaatcaaacctctcttggaacatctgcgcctgagcttgggaagcttgac	386
Db	86684	TAGATGTCCACACAACTACATCCGAGATATTCCTGCTCAGTTGGGAACTTGA	86625
QY	387	attacagctatctagagctctcaaaaacacatcccaaggaactatacttcggaactg	446
Db	86624	ATCTAGACACAGCTTGATTGTGAGCAGAAACACCTTCACAGAAACATACCGTGACTTTG	86565
QY	447	gaactctgaagaatctcatcagcttgatctgctacaaacaaatcttaagaagatgctc	506
Db	86564	GGAACTTAAGCTACTTTAAACAAGCTGATCTCCACAAACAATCTCCTCACTGGTCAATCC	86505
QY	507	ccactcttttgggaatctgaagctctcgtgctcttttcaggcttaatgacacagatga	566
Db	86504	CAAAATCTATCAAGATCTATCAAAAAGCTTTACACTTCTGATCTTGACTTCAACAGCTTAT	86445
QY	567	ccggtccatccctctagagcactcacggaatcccaagccttaagttgtgagctctca	626
Db	86444	CCGGTGAATATCCACAGAAAGACTCGGTCAAGTCACAAAGCTTAACTTAAACCTCGACTTAA	86385
QY	627	gcaa 630	
Db	86384	GCTA 86381	

RESULT 15
ATAC011620
LOCUS ATAC011620 100887 bp DNA PLN 18-JAN-2000
DEFINITION Arabidopsis thaliana chromosome III BAC F18C1 genomic sequence,
complete sequence.
AC011620
AC011620.6 GI:6714437
HTG.
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 100887)
Lin,X., Kaul,S., Town,C.D., Benito,M., Greasy,T.H., Haas,B., Wu,D.,
Ronning,C.M., Koo,H., Fujii,C.Y., Uterback,T.R., Barnstead,M.E.,
Bowman,C.E., White,O., Niernann,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome III BAC F18C1 genomic sequence
Unpublished
2 (bases 1 to 100887)
Lin,X. and Kaul,S.
Direct Submission
Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, xlinetlgr.org
3 (bases 1 to 100887)
Lin,X.
Direct Submission
Submitted (18-JAN-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jan 18, 2000 this sequence version replaced gi:671515.
Address all correspondence to: atetlgr.org

COMMENT

BAC clone F18C1 is from Arabidopsis chromosome III and is near the
molecular marker ml172.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL (available by anonymous ftp
from athurh.chem.ornl.gov), GeneFinder (Phil Green, University of
Washington), GenScan (Chris Burge,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:52:36 ; Search time 446.21 Seconds
(without alignments)
550.051 Million cell updates/sec

Title: US-09-180-798-24

Perfect score: 981
Sequence: 1 agtggagtaattagttg.....tcgcctaattaccacact 981

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 segs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	981	100.0	981	1 V06587	Arabidopsis thaliana
2	924	94.2	1106	1 V06587	Arabidopsis thaliana
3	841.6	85.8	1063	1 V06590	Arabidopsis thaliana
4	742	75.6	788	1 V06588	Arabidopsis thaliana
5	356.2	36.3	894	1 V06589	Arabidopsis thaliana
6	210.2	21.4	2089	1 V06591	Arabidopsis thaliana
7	101.8	10.4	1814	1 V06571	Daucus carota SERK
8	75	7.6	5940	1 X23526	O. longistaminata X
9	74	7.5	6695	1 V06570	Daucus carota SERK
10	73.8	7.5	3176	1 T62124	Arabidopsis thaliana
11	72.4	7.4	4081	1 V06585	Arabidopsis thaliana
12	71.8	7.3	9424	1 X23525	O. sativa Xa21 gen
13	68.4	7.0	3573	1 T06307	Partial tomato pat
14	68.4	7.0	6471	1 T06306	Tomato pathogen re
15	67.6	6.9	8416	1 X23523	O. longistaminata
16	67.4	6.9	3979	1 V14518	CF-5 pathogen resi
17	67.4	6.8	3979	1 V14519	CF-5 pathogen resi
18	67.4	6.9	4123	1 V14523	CF-5 pathogen resi
19	67	6.8	3921	1 T31300	Rice Xa21 disease
20	67	6.8	13340	1 X23522	O. longistaminata
21	67	6.8	19639	1 X23524	O. longistaminata
22	66.2	6.7	3541	1 V14522	CF-5 pathogen resi
23	65.6	6.7	3045	1 X23531	Maize Xa21 gene DT
24	63.8	6.5	7204	1 X23527	O. longistaminata X
25	62.6	6.4	3293	1 X23532	Tomato Xa21 clone
26	60.2	6.1	2075	1 T49435	Tomato polygalactu
27	60.2	6.1	6256	1 T31299	Rice Xa21 disease
28	59.6	6.1	2192	1 X23530	Maize Xa21 gene DT
29	56.8	5.8	1058	1 T49434	Pear polygalactu
30	54.6	5.6	5733	1 X00477	Arabidopsis thaliana
31	54.2	5.5	1554	1 T31307	Tomato RRR gene cl
32	51.6	5.3	4104	1 X07356	Arabidopsis thaliana
33	51.6	5.3	3842	1 X23533	Tomato Xa21 clone
34	49.8	5.1	3050	1 T06309	Tomato pathogen re

35	49.8	5.1	3089	1 T47877	Tomato pathogen re
36	48.4	4.9	2880	1 G91450	Tomato Cf-9 cDNA.
37	48.4	4.9	2880	1 T06304	Tomato pathogen re
38	48.4	4.9	3905	1 G91449	Tomato Cf-9 gene.
39	48.4	4.9	3905	1 T06303	Tomato pathogen re
40	41.4	4.2	1117	1 Q42595	Sequence encoding
41	41.4	4.2	2917	1 T86755	Rasperry drul gen
42	40.8	4.2	792	1 Q42596	Sequence of varian
43	37.6	3.8	2755	1 X52230	Protein PRO266 cDN
44	36.4	3.7	457	1 X21421	5' and 3' regulatio
45	36.4	3.7	1448	1 X21416	Selectable marker

ALIGNMENTS

RESULT	1	
ID	V06587	standard; cDNA to mRNA; 981 BP.
AC	V06587;	
DT	03-AUG-1998	(first entry)
DE	Arabidopsis thaliana SERK LRR homologous EST clone.	
KW	receptor kinase; apomictis; apomictic; seeds; production; embryos;	
KW	plant breeding; leucine-rich repeat; ss.	
OS	Arabidopsis thaliana.	
FH	Key	Location/Qualifiers
FT	CDS	104..760
FT		/tag= a
FT		/note= "shows high homology to SERK"
PN	W09743427-A1.	
PD	20-NOV-1997.	
PF	13-NOV-1997; E02443.	
PR	14-MAY-1996; GB-010044.	
PA	(NOVS) NOVARTIS AG.	
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;	
DR	WPI; 98-086529/08.	
DR	P-PSDB; W47019.	
PT	Production of apomictic seeds - useful in plant breeding	
PS	Claim 28: Pages 75-77; 123pp; English.	
CC	The sequence is that of an EST clone showing high homology to	
CC	SERK LRR (leucine-rich repeat) sequences. 180 G; 279 T;	
SQ	Sequence 981 BP; 286 A; 236 C; 279 T;	
Query Match	100.0%; Score 981; DB 1; Length 981;	
Best local Similarity	100.0%; Pred. No. 2, 5e-262;	
Matches	981; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	agtggagtaattagttgcttcctcctcttgctcagaaatttccttactctca 60
DB	1	AGTGTGAGTAAATTAGTTGCTTCCTCCCTTGTTCAGAAATTTCCCTTACTCTCA 60
QY	61	aattcctttcgattccctcctccttaaacctcgaagctcagtggtctgaaacta 120
DB	61	AAATTCCTTTTCGATTCCTCCTCTTAAACCTCGAAGCTCAATGAGCTCGAACTA 120
QY	121	tcggtggagctcttcgcagcttcgttaacccaactagattgattcaccctgtcga 180
DB	121	TCGGTGGAGCTCTTCGAGCTTCGTTAACCTTAACCTTAGCTTGAATTCACCTGTGA 180
QY	181	agcaaatccgaaggagatgctctcctccttcgacgcttcgacgagttggaagatccagacaa 240
DB	181	AGCAAACTCCGAAGGAGATGCTCTCTACCTCTTCCCGAGATTGACAGATCCAGACCA 240
QY	241	tgctccagagcttggaatcccaactcttgtaactcttgtaactggttcacatgtaacctg 300
DB	241	TGTCCTCCAGAGCTGGAGATCCCAACTCTGTATATCCTTGATACCTGTTCCATGTCACCTG 300
QY	301	taaccaagaacacccgctcaactcgctgtggaattgggaaattcaaacctctctgacatct 360
DB	301	TAAACCAAGACCAACCGCTCACTCGTGTGATTTGGAAATTCATCACTCTGACATCT 360
QY	361	tcgcctgaagcttggaagcttgtaacattacagatctagagctcctcaaaaacaacat 420

```
|||||
Db 361 TGGCGTGAACCTTGGAGACCTTGAACATTACAGTATCTAGAGCTCTTACAAAAAACAACCT 420
Qy 421 ccaaggacataccttcggaacttgaaactcgaagaaatccatcaactgaacttgatctgta 480
Db 421 CCAAGGAACATACCTTCGGAATCTGGAATCTGGAATCTCATCAGCTTGATGATCTGTA 480
Qy 481 caaacaacatctcaaggagtagttcccaactcttgggaaaaattgaagtcctgctt 540
Db 481 CAACAACAACTTACAGAGGATAGTCCCACTTTGGGAAAAATTGAATCTCTGCTCTT 540
Qy 541 tttaagcgttaaatgagaaacggaattgacggtcccaatccctagagacactgaagcaatccc 600
Db 541 TTTACGGCTTAATGACACCGATGACCGGTCCAACTCCTAGAGCACTACCGCAATCCC 600
Qy 601 aagccttaagttgtgagcgtctcaagcaatgattgtgtggaacaatcccaacaacg 660
Db 601 AAGCCTTAAGTGTGAGCTCTCAAGCAATGATTTGTGGMAACAATCCCAACAACGG 660
Qy 661 acccttggccacattccctttaagaactttggaacaacccgagattggagggaccgga 720
Db 661 ACCCTTGGCTCACATTCCTTTACGAACCTTGGAACAACCCGAGATTGGAGGACCGGA 720
Qy 721 attactcgtctgcaagcagcaactaactgacactgacactgaacaactggcaaacctgaa 780
Db 721 ATTACTCGCTGCTGCAAGCTAGACACTAATCTGACACTGAAACAATGGCAAAACCTGAA 780
Qy 781 aatgaagaattggggggtgacctgtgaagaacacacacacacacacacacacacacac 840
Db 781 AATGAAGAATTGGGGGTGACCTGTGAAGAACCTTCACACACTTTATCAAAATATCAACAC 840
Qy 841 tattaatgaataagatatataatctagtataaacaacaaataatgaagaatcgatcgta 900
Db 841 TATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 900
Qy 901 atatatcctggtcctcaattgagaactcgaaggtcgtatgtaaaatttcbaatcgatt 960
Db 901 ATATCATCTGCTGCTCAATGAGAACTTCGAGCTGTATATATAAATTCTAATTCGAAAT 960
Qy 961 ttggccttaattactacact 981
Db 961 TTGCGCTTAATTAATCTACACT 981

RESULT 2
ID V06586 standard: cDNA to mRNA; 1106 BP.
AC V06586;
DT 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS
FT FT 142..798
FT FT /*tag= a
FT FT /note= "shows high homology to SERK"
PD W09743427-A1.
PD 20-NOV-1997.
PD 13-MAY-1997: E02443.
PR 14-MAY-1996: GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WP1: 98-086529/08.
DR P-PSDB: W47018.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28: Pages 71-73: 123pp: English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;
```

```
Best Local Similarity 98.8%; Pred. No. 1,6e-246;
Matches 952; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
Qy 13 ttagttgcttcctcctcttgcttgctgagaaatttcccttactctcaatctcttctg 72
Db 51 TAATTTGCTTCCCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 110
Qy 73 atttccctccttaaaactccgaaagctcaaatggtcttgaaactatcggttgagct 132
Db 111 ATTTCCTCTCTTAACCTCGAAAGCTCACATGGGTCTCGAAATATCGGTGGAGCT 170
Qy 133 ctgcgagcttgcttaacccctaaccttaagcttgatctgacacgtgctgaaagcaactcga 192
Db 171 CTTCGAGCTTCCTGTTAACCCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCT 230
Qy 193 agagatgctctctcagctcttcgagagttgacagatccagacatgctctcagag 252
Db 231 AGGAGATGCTCTACCTCTTCGCGGAGTTGACAGATCCAGACCATGTCTCCAGAG 290
Qy 253 ctggagatccacactctgttaactctgttactgtgctgacatgctgacactgtaacgaagaa 312
Db 291 CTGGGATCCAACTCTTTATCTCTTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350
Qy 313 ccggtcactcgtgtggaattgggaaattcaaacctctctgacacactctgagct 372
Db 351 CCGGCTCACTGCTGTGATTTGGGAATTCAAACCTCTGACATCTTGCGCTGAGCT 410
Qy 373 tgggaagcttgaagaattacaatctcagagctctcaaaaaaacaataccaagaaactat 432
Db 411 TGGGAAGCTTGAAATTTAAGATTTCTAGAGCTCTCAAAAAAACAATCCAAAGAACTAT 470
Qy 433 accctcgaacttgaatctgaagaatccatccacacttgatctgtatcaacaacacatct 492
Db 471 ACCTTCGAATTTGGAATCTGAAGATCTCATCAGTTGATCTGTACACAAACATCT 530
Qy 493 taagsgatagttccacactcttgaggaaattgaagctctgctcttcttcttcttcttct 552
Db 531 TACAGGATGATCTCCACTTCTTGGAATTTGAAGTCTCTGCTCTTCTTCTTCTTCTT 590
Qy 553 tgacaacggaattgacggtcgaatcccttagagacatccgaagcaactcccaagcc-ttaaag 611
Db 591 TGACACCGATTGACCGGTCCAAAT-CCTAGACATCTACCGCAATCCCAAGCCTTTAAAG 649
Qy 612 ttgttgacgtctcaagaatgattgtgtggaacaaatcccaacaacggaacacttgctc 671
Db 650 TTGTTGAGCTCTCAAGCAATGATTTGTGGACAAATCCCAACAAACGACCTTTGCTC 709
Qy 672 acattcctttcaagaactttgagaaacacccgagattggagggacgggaattcctggtc 731
Db 710 ACATTCCTTTACAGAACTTTGAGAACACCCGAGATTGGAGGACCGGAATTTACTGGTTC 769
Qy 732 ttgcaagctagacaactcgtcacctgaaacaactgycacaacactgaaacactgaaatgaagaat 791
Db 770 TTGCAAGTATGACACTAATCTGACCTGAAACAACATCTGCAAAACCTGAAATAAGAAAT 829
Qy 792 ggggggtgacctgtgaagaacacttccacacttatcaaatatccatctatattat 851
Db 830 GGGGGGTGACCTTTGAGAACACTTCAACCTTATATCACTTATATCACTTATATATAT 889
Qy 852 aagtatataatgagtaaaacaaacaaataatgaagaatgaagaatgaagaatgaagaat 911
Db 890 AAGTATATATGTATGTAATTAACAAACAAATGAAGATGCAATCGTATATATCATCTGG 949
Qy 912 tctcaattgaagaactcgaaggtcgtatgtlaaaatttctaaatgagatttgcctaaat 971
Db 950 TCTCATTTGAGAACTTGAGAGTCTGTATGTAATTTCTTAATGCAATTTCTGCTACTG 1009
Qy 972 tact 975
Db 1010 TAAT 1013

RESULT 3
```


Y065590	ID	V065590	standard; cDNA to mRNA; 1063 BP.
AC	AC	V065590;	
DT	03-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK LRR homologous EST clone.		
KW	receptor kinases; apomixis; apomictic; seeds; production; embryos;		
KW	plant breeding; leucine-rich repeat; ss.		
OS	Arabidopsis thaliana.		
FH	Key	Location/Qualifiers	
FT	CDS	106..762	
FT		/tag="a	
FT		/note="shows high homology to SERK"	
PN	NO97A3427-A1.		
PD	20-NOV-1997.		
PF	13-MAY-1997; E02443.		
PR	14-MAY-1996; GB-010044.		
PA	(NOVS) NOVARTIS AG.		
PI	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;		
PI	WPI: 98-086529/08.		
DR	P-PSDB: M47022.		
DR			
PT	Production of apomictic seeds - useful in plant breeding		
PS	claim 26; Pages 86-88; 123pp; English.		
CC	The sequence is that of an EST clone showing high homology to		
CC	SERK LRR (Jaucune-rich repeat) sequences.		
CC	Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;		

Query Match	85.8%	Score 84.1	8: DB.1	Length 1063:
Best Local Similarity	96.1%	Pred. NO. 8.9e-224:		
Matches 885:	Conservative	0:	Mismatches 32:	Indels 4: Gaps 2:
QY	ctcaaatctcttcgattccctctctcaaacctccgaagctcaatgcygclctogaa	116		
Db	CTCAAAATCCTAATTAATTAATCTCTCTCTCGACCTCCCATAGCTCAATGAGCGCTCGAA	118		
OY	actatcggtggagagccttcgcagctctgttaaccttaaccttggctttgatcaactg	176		
Db	ACTATCGGTGGAGAGCTTCTTCGACCTCTGTAACTTAATCCTTAACCTTAGCTTTGATCACTGG	178		
OY	tcgaagcaaacctcgaagagabgcctctcaagccttcgcgcgagatttcagacatcag	236		
Db	TCGAGCAAACTCCGAAGAGAGTGCCTTTAGCGCTTCGCCGGAGATTAAAGATCCGG	238		
OY	accatgcccacagagctggagatcccaacctgttaacctcttgaactggttccatgata	296		
Db	ACCATGCTTCCAGAGAGCTGGAGATCCAACTCTGTAAATCCTTGACCTGGTTCATGTCA	298		
OY	cctgtlaaccaagaacacgcgctacccgctgctgattgtgaatttggaaattcaaacctctc	356		
Db	CCTGTACCAAGAACACCGCGCTACCGTGTGATTTGGGGAATTCAAACTCTCTGTGAC	358		
OY	atcttgagccttgagcttggagaagcttgaacattcaatctgaatcgtagctctcaaaaaa	416		
Db	ATCTTGAGCCTTGAGCTTGGAGAGCTTGAACATTTCAGATCTGAGCTCTACAAAAACA	418		
OY	acatccaaggaaactataccttcgcgaacttggaaatctgaagatcacaatcagcttgatc	476		
Db	ACATCCAGAGAACTATACCTTCGAACTGGAAATCTGAAGATCTCATCAGCTTGATC	478		
OY	tgtaacaacaacatcttaacaggaatgcttcccaactcttttgggaaattgaagctcttg	536		
Db	TGTACAAACAACATCTTACAGGAGATGTTCCCACTCTTTGGGAAATAATGAAGTCTGTGG	538		
OY	tcttttaacgcttaabgaacacggatttgcacggtccatccctctagacatcaacgycaa	596		
Db	TCTTTTACGGCTTAATGAACAACCGATTGACGGGCGCAATCCCTTGAGCACTCACTGCA	598		
OY	tcccaagcccttaagatgttgagctctcaagaagatatttcttggagaacaatcccaaca	656		
Db	TCCCAAGCCTTAAAGTGTGATGCTCTCAAGCAATGATTTGTGTGGAACAATCCCAACAA	658		
OY	acggaccccttgctcacatctcttcaagaacttggagaacaacccgagatttgaagggac	716		

Db	659	ACGACCTTTTGGCTCACACTTCCTTTACAGAACTTTGAGAACCAACCCAGAGTTGGAGGGAC	718
Qy	717	cgggaatbaactcgggctctgcgaagctacgacacactatgcacacttgaaacacacttgcgaacacc	776
Db	719	CGGAATTACTCGGTCTTTCGAAGCTACGACACTAATGCGACCTGAAAAAATTGGCAAAACC	778
Qy	777	tgaaatggaagaatctgggggggtgcacctgtcaagaacacttcacacactttatcaaatatca	836
Db	779	TGAAATGGAAGAATTGGGGGGGTGAACCTTGTAGACACTTCACACTTTATCAAAATATCA	838
Qy	837	catctattatgtcaataagtatatatatagtgtgaataacaaataaataatgaagaatcgaatc	896
Db	839	CATCTACTATGTAATADAGATATATATATAGTGTCCAA--AAAAAAATAGAAATCGAATC	898
Qy	897	ggtataatcatctcgtgcctaaatggaagaacttcgaagtc--tgtatgtaaaatcttctaatt	954
Db	897	AGTAATATCATCGTGGGTCCTCAATTGGAACCTTGGAGTCGTGTGTATGTAATATTTCTAAAT	956
Qy	955	ggcgaatttgcgcctaaataact 975	
Db	957	GGCAGCTTTCGCGTACTGTAAT 977	

ID	Accession	Location/Qualifiers
AC	V06588	standard; cDNA to mRNA; 788 bp.
DT	03-AUG-1998	(first entry)
DE	Arabidopsis thaliana SERK LRR homologous EST clone.	
KM	receptor kinases; apomictic; seeds; production; embryos;	
KW	plant breeding; leucine-rich repeat; ss.	
OS	Arabidopsis thaliana.	
FT	Key	Location/Qualifiers
FT	CDS	2..664
FT	/*tag= a	/note= "Shows high homology to SERK"
PD	W09743427-A1.	
PD	20-NOV-1997.	
PF	13-MAY-1997; E02443.	
PR	14-MAY-1996; GB-010044.	
PA	(NOVS) NOVARTIS AG.	
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;	
DR	WPI: 98-085529/08.	
DR	P-FSDS; W47020.	
PT	Production of apomictic seeds - useful in plant breeding	
PS	Claim 28; Pages 79-80; 123pp; English.	
CC	The sequence is that of an EST clone showing high homology to	
CC	SERK LRR (leucine-rich repeat) sequences.	
SO	Sequence 788 bp; 234 A; 191 C; 156 G; 207 T;	

DT 17-JUN-1999 (first entry)
 DE O. longistaminata Xa21 gene family member A2 DNA.
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice;
 KM plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
 OS Oryza longistaminata.
 PN MO9909151-A2.
 PD 25-FEB-1998.
 PF 17-JUL-1998; U14841.
 PR 13-AUG-1997; US-910386.
 PA (REGC) UNIV CALIFORNIA.
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,
 PI Wang G:
 PI WPI: 99-204431/17.
 PT New RRR polynucleotides and nucleic acid constructs - used for
 PT generating transgenic plants resistant to Xanthomonas
 PS Claim 1; Page 52-53; 67pp; English.
 CC This invention describes a method for conferring disease resistance in
 CC plants. The invention describes the use of novel genes and proteins
 CC belonging to the Oryza longistaminata and Oryza sativa receptor
 CC kinase-like protein (RRR) Xa21 multigene family. Such genes from
 CC cassava, maize and tomato are also described. The genes and proteins can
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in
 CC rice or tomato.
 SQ Sequence 5940 BP; 1570 A; 1200 C; 1188 G; 1982 T;

Query Match 7.6%; Score 75; DB 1; Length 5940;

Best Local Similarity 52.4%; Pred. No. 3.3e-11;

Matches 165; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 341 tcaacctctctgcatctgagctgagctggaagctggaacattatgacatctca 400
 DB 2406 TCGGACTGTCGGGATCATCTCGCGCTGGGCAACCTCTCTCTCTGAGACGCTG 2465
 QY 401 gagctctcaaaaacaacatccaagaactatactccgaactggaatctgagaat 460
 DB 2466 GACCTAGCGCACACCACTGTCCGGCAAGATACCTTAGGAAGCTCAGCAGTCTCAGCAGG 2525
 QY 461 ctatcagctgagctgctgtaacacacacatctacagagatagttccactctcttgga 520
 DB 2526 CTCCAACAACCTGATCTGATTTCAACACGCTATCGGCTGAGATTCACGCTCTTGAGGC 2585
 QY 521 aaattgaagctctgctcttttaagcgttaatgacacacgaattgacggtccaatccct 580
 DB 2586 AACTTAACCACTCTCGGCTGTTGAGCTGACTAACAATACACTGCTGAGCAATCCCT 2645
 QY 581 agagcactcagcgaatcccaagccttaagttgttgagcgttccaagcaatgttgt 640
 DB 2646 TCATCTCTGGGCAAACTCACAGCTCATGATCTTGCACTGGCTGAAATAATATGCTGCT 2705
 QY 641 ggaacaatcccaaca 655
 DB 2706 AGTTCCATCCCTTCA 2720

RESULT 9

V06570 standard; DNA; 6695 BP.

AC V06570;
 DT 03-AUG-1998 (first entry)
 DE Daucus carota SERK gene.
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
 OS Daucus carota.
 FH Key
 FT CDS Location/Qualifiers
 FT 3696..6620
 FT /*tag= a
 FT /note= "contains introns"
 FT 3731..3802
 FT /*tag= b
 FT /number= 1
 FT 3851..3979
 FT /*tag= c
 FT Intron
 FT Intron

FT /number= 2
 FT Intron .4124..4211
 FT /*tag= d
 FT Intron /number= 3
 FT .4284..4357
 FT /*tag= e
 FT Intron /number= 4
 FT .4430..4528
 FT /*tag= f
 FT Intron /number= 5
 FT 4642..4757
 FT /*tag= g
 FT Intron /number= 6
 FT 4890..4967
 FT /*tag= h
 FT Intron /number= 7
 FT 5295..5803
 FT /*tag= i
 FT Intron /number= 8
 FT 6197..6339
 FT /*tag= j
 FT /number= 9
 PN WO9743427-A1.
 PD 20-NOV-1997.
 PE 13-MAY-1997; E02443.
 PR 14-MAY-1996; GB-010044.
 PA (NOVS) NOVARTIS AG.
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
 PI WPI: 98-086529/08.
 DR P-PSDB; W47013.
 PT Production of apomictic seeds - useful in plant breeding
 PS Claim 21; Pages 40-46; 123pp; English.
 CC The sequence is that encoding SERK, a putative receptor kinase.
 CC It may be used as part of a method of producing apomictic seeds
 CC comprising: (a) transforming plant material with a nucleotide
 CC sequence encoding a protein which in active form in a cell or
 CC cell membrane renders the cell embryogenic; (b) regenerating
 CC the transformed material into plants or carpel-containing
 CC plant parts; and (c) expressing the sequence in the vicinity
 CC of the embryo sac. The apomictic seeds and embryos thus produced
 CC can be developed into plant progeny. This is useful in plant
 CC breeding programs. Controllable and reproducible apomixis provides
 CC many advantages in plant improvement and cultivar development in
 CC the case that sexual plants are available as crosses with the
 CC apomictic plant. Apomixis provides for true-breeding, seed
 CC propagated hybrids and could shorten and simplify the breeding
 CC process so that selfing and progeny testing to produce and/or
 CC stabilise a desirable gene combination could be eliminated.
 CC Apomixis allows plant breeders to develop cultivars with
 CC specific stable traits for such characteristics as height,
 CC seed and forage quality and maturity.
 SQ Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match 7.5%; Score 74; DB 1; Length 6695;

Best Local Similarity 71.0%; Pred. No. 6.6e-11;

Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 193 agagatgctctctgactctctcgcgaggttgagacatccagacatgctccagag 252
 DB 956 AGCGATGCACTTACACAACTTACGAACTTACGCAAGATCCCAACAAATGCTCTGAGAG 1015
 QY 253 ctggatccaactctgttaatcctctgttaccctggttcacatgtaacctgaacaaagaca 312
 DB 1016 CTGGGATCCAAACCTTGTGAACCTTGACATGTTTCATGTGACATGAACAATGAANA 1075
 QY 313 ccgcgtcaactcgtgtgga 330
 DB 1076 CAGTGTATAGAGTGA 1093

RESULT 10

T62124

Qy 313 ccgcgcactcgtgtga 330
1 | | | | |
Db 1914 CAGTGTCAATAGAGTSTA 1931

RESULT 12

X23525
X23525 standard; DNA; 9424 BP.

AC X23525;
DE 17-JUN-1999 (first entry)
KW O. sativa Xa21 gene family member E DNA.
KW Xa21: receptor kinase-like protein; multigene family; Rf; rice;
KW Plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
OS Oryza sativa.
PN WO9909151-A2.
PD 25-FEB-1999.
PF 17-JUL-1998; 014841.
PR 13-NOV-1997; US-910386.
PA (REGC) UNIV CALIFORNIA.
PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V, Wang G;
DR WPI: 99-204431/17.
DR P-PSDB: W93598.
PT New RRK polynucleotides and nucleic acid constructs - used for generating transgenic plants resistant to Xanthomonas
PS Claim 1; Page 48-50; 67pp; English.
CC This invention describes a method for conferring disease resistance in plants. The invention describes the use of novel genes and proteins belonging to the Oryza longistaminata and Oryza sativa receptor kinase-like protein (RRK) Xa21 multigene family. Such genes from cassava, maize and tomato are also described. The genes and proteins can be used for enhancing resistance to Xanthomonas in a plant, preferably in rice or tomato.
SQ Sequence 9424 BP; 2681 A; 2138 C; 2053 G; 2552 T;

Query Match 7.3%; Score 71.8; DB 1; Length 9424;

Best Local Similarity 51.7%; Pred. No. 3.1e-10; Mismatches 152; Indels 0; Gaps 0;

Qy 341 tcaaacctcctcgacatcttcgctgagcttggaagcttgaacattacagatctca 400
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 3080 TCCAACTGGCGGGATCATCTCGCTGCGTGGCAACTTACCTTCACAGAGCGTG 3139
Qy 401 gagctctcaaaaacacatccaaggaactatccttcggaacttgaagaaat 460
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 3140 CAACCTCAGCAGCAACACCTGTCGCGCAAGATACCCAGAGAGCTCAGCCCTCAGCAGG 3199
Qy 461 ctcatcagcttgatctgtacaaacacacattacagaggaatgtcccaactcttgggga 520
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 3200 CTCACGACAACTGACTGAATTTTACAGGCTATCGGGTGAAGATTCAGTGTGTTGGCG 3259
Qy 521 aaattgaagtcctggtctttttacagcttaagacaacagatgacggttcacatccct 580
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 3260 AATCTAACAGTCTCTCGGTTCTTGAAGCTACATCAATACATGTCGAGCAATCCCT 3319
Qy 581 agagcactcaagcgaatcccaagccttaagttgttgcgtcccaagcaatgttgtg 640
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 3320 TCACCTCTGGGCAACGTCACAGGTCATGATGTGCTGCTGCTGCTGCTGCTGCTGCT 3379
Qy 641 ggaacaatcccaaca 655
1 | | | | | 1 | | | | |
Db 3380 GGTTCATCCCATCA 3394

RESULT 13

T06307
T06307 standard; cDNA; 3573 BP.

AC T06307;
DE 14-APR-1996 (first entry)
KW Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.
KW Pathogen resistant; Cf-2.2; tomato; C.fulvum; Avr 4; Avr 9; fungal;

KW leaf mould; variegation; ss.
OS Lycopersicon esculentum.
PN WO9531564-A2.
PD 23-NOV-1995.

PF 11-MAY-1995; G01075.
PR 11-MAY-1994; GB-009394.
PR 23-DEC-1994; NO-G02812.
PR 31-MAR-1995; GB-006658.
PR 07-APR-1995; GB-007232.
PA (GATS-) GATSBY CHARITABLE FOUND.
PI Hammond-Kosack KE, Jones DA, Jones JDG;
DR WPI: 96-010949/01.
DR P-PSDB: R85399.
PT Increasing plant pathogen resistance by induction of variegation - may lead to acquired resistance to a broad range of pathogens.
PS Claim 9; Page 85-87; 131pp; English.
CC T06307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone. In a new method this gene is expressed highly in genetic constructs CC which may be used to impart a broad range of pathogen resistance, by induction of variegation, to transgenic plants (or parts or propagules of plants) containing such constructs. Cf-2.2 imparts resistance to the disease caused by the leaf mould fungal pathogen Cladosporium fulvum. C.fulvum contains avirulence (Avr) genes that confer recognition by CC plants containing Cf-genes, leading to the activation of host defence mechanisms to attack the disease.
SQ Sequence 3573 BP; 1032 A; 654 C; 664 G; 1223 T;

Query Match 7.0%; Score 68.4; DB 1; Length 3573;

Best Local Similarity 52.9%; Pred. No. 1.8e-09; Mismatches 147; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 311 aaccgcgtcactcgtgtgagcttggtggaattcaacctcttgagacttcgctcgtgag 370
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 995 AAAACCTGTCTAGCTGAATCTGTATATATCAGCTTTCGCTTATTCCTGCTTCA 1054
Qy 371 ctgggaagcttgacattacagatctagagctctacaacaaacacatccaaggaact 430
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 1055 TTGGGGAATCTGAACAACTTGTATGTATTTTACATATACAGCTTTCGCTCT 1114
Qy 431 ataccctcgaacttggaactctgaagaatctcatcagcttgatctgtctgtacacaaat 490
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 1115 ATTCCTGCTTATTTGGGGAATCTGAACAACTTGTATGTATTTTACATATACAG 1174
Qy 491 cttaagagtagttcccaactcttgggaatgaatgaatgagtcgtctttttagagctt 550
1 | | | | | 1 | | | | | 1 | | | | | 1 | | | | |
Db 1175 CTTCGTGCTTATTCCTGCTTATTTGGGGAATCTGAACAACTTGTATGTATTC 1234
Qy 551 aatgacaacgattgaccggtcccaatcccttagaggaact 588
1 | | | | | 1 | | | | | 1 | | | | |
Db 1235 TACAAATATCAGCTTTCGCTTATTCCTGAAGAAAT 1272

RESULT 14

T06306
T06306 standard; DNA; 6471 BP.

AC T06306;
DE 14-APR-1996 (first entry)
KW Tomato pathogen resistance gene Cf-2.1.
KW Pathogen resistant; Cf-2.1; tomato; C.fulvum; Avr 4; Avr 9; fungal;
KW Leaf mould; variegation; ds.
OS Lycopersicon esculentum.

OS Lycopersicon esculentum.
FH Key location/Qualifiers
FT 5'utr 1..1676
FT 1677..5014

cds

FT 1677..5014
FT /*tag= a
FT /*tag= b
FT /*product= immature_Cf-2.1_protein

signal_peptide

FT 1677..1745
FT /*tag= c
FT mat_peptide 1746..5011

3'utr

FT 5015..6471

FT W0931564-A2. /tag= e
PN 23-NOV-1995.
PD 11-MAY-1995; G01075.
PR 11-MAY-1994; GB-009394.
PR 23-DEC-1994; WO-602812.
PR 31-MAR-1995; GB-006658.
PR 07-APR-1995; GB-007232.
PA (GATS-) GATSBY CHARITABLE FOUND.
PI Hammond-Kosack KE, Jones DA, Jones JDG;
DR WPI: 96-010949/01.
PT P-PSDB; R85298.
PS Increasing plant pathogen resistance by induction of variegation -
PT may lead to acquired resistance to a broad range of pathogens.
PS Claim 9; Page 80-83; 131pp; English.
CC This gene is expressed highly in genetic constructs which may be used
CC to impart a broad range of pathogen resistance, by induction of
CC variegation, to transgenic plants (or parts or propagules of plants)
CC containing such constructs. Cf-2.1 imparts resistance to the disease
CC caused by the leaf mould fungal pathogen Cladosporium fulvum.
CC C. fulvum contains avirulence (Avr) genes that confer recognition by
CC plants containing Cf-genes, leading to the activation of host
CC defence mechanisms to attack the disease.
SQ Sequence 6471 BP; 2073 A; 1106 C; 1122 G; 2170 T;

Query Match 7.0%; Score 68.4; DB 1; Length 6471;
Best Local Similarity 52.9%; Pred.No.2.3e-09;
Matches 147; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 311 aaccgcgcacgcctgctgagcttgggaatcaaacctctcgagacatctgcgcctgag 370
DB 2679 AAAAATTGCTGAGGTGATCTGTAATATCATCTTCTGGCTCTATCTCTGCTCA 2738
QY 371 ctgggaagcttgaaacttcaagctatctagagctctacaacaaacactccaaggaact 430
DB 2739 TTGGGGATCTGAACAACCTGTATGTATCTTAATAATCAACGCTTTCTGGCTCT 2798
QY 431 ataccctccgaacttggaatctgaagaatcctcatcagcttgatctgtacaacacaaat 490
DB 2799 ATTCTGCTTCATTTGGGGAATCTGAACAACCTGTCTATGTGATCTTACAAATATG 2858
QY 491 cttaacaggagtagtcccaattcttgggaaatgaagctcctgctcttttaagcctt 550
DB 2859 CTTCTGCTCTATCTCTGCTCATTTGGGGAATCTGAACAACCTGTCTAGGTTGATCTC 2918
QY 551 aatgacacagattgacgcgtccatccctagagcaact 588
DB 2919 TACAATAATCAGCTTCTGCTCTATCTCTGAGAAAT 2956

RESULT 15
X23523 X23523 standard; DNA; 8416 BP.
AC X23523;
DT 17-JUN-1999 (first entry)
DE O. longistaminata Xa21 gene family member A1 DNA.
KW Xa21; receptor kinase-like protein; multigene family; RKK; rice;
OS plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
PN WO9909151-A2.
PD 25-FEB-1999.
PE 17-JUL-1998; U14841.
PR 13-AUG-1997; US-910386.
PA (REGC) UNIV CALIFORNIA.
PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,
PI Wang G;
DR WPI: 99-204431/17.
DR P-PSDB; W93596.
PT New RKK polynucleotides and nucleic acid constructs - used for
PT generating transgenic plants resistant to Xanthomonas
PS Claim 1; Page 37-39; 67pp; English.

CC This invention describes a method for conferring disease resistance in
CC plants. The invention describes the use of novel genes and proteins
CC belonging to the Oryza longistaminata and Oryza sativa receptor
CC kinase-like protein (RKK) Xa21 multigene family. Such genes from
CC cassava, maize and tomato are also described. The genes and proteins can
CC be used for enhancing resistance to Xanthomonas in a plant, preferably in
CC rice or tomato.
SQ Sequence 8416 BP; 2220 A; 1984 C; 1707 G; 2505 T;

Query Match 6.9%; Score 67.6; DB 1; Length 8416;
Best Local Similarity 51.0%; Pred.No.4.3e-09;
Matches 160; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 341 tcaacctctcgacatcttgccttgaagcttgggaagcttgaaattacagatcata 400
DB 5026 TCCAACTCTGACGGGATCTCTCGCAATCGCTGGGCAACCTATCTCTCAGAGCCTG 5085
QY 401 gagctctacaacaacacatccaaggaactatcctccgaacttggaattgaagaat 460
DB 5086 CAACTCAGCAACACACCTCTCCGCAAGATACCCAGAGCTCAGCCGCTCAGCAGG 5145
QY 461 ctatcagcttgatctgtacaacaacaactcttaagagatgtccactcttggga 520
DB 5146 CTCACGACGCTGTAATTTCAACAGCCATGCGGTGAGATTCACGCTGTTGGGC 5205
QY 521 aaattgaagctctgctctttaaagcttaaggaacacagattgacgcgtcaatccct 580
DB 5206 AATCTAACAGTCTCTCAATCTTGAAGCTTAACAAATACAGTCTGAGTTTATCTCT 5265
QY 581 agagcactacggcaatcccaagccttaagttgttgacgtctcaagaatgatttgt 640
DB 5266 TCATCCCTGGCAAGCTCACCGGCTCTATCTTGACAGCTGGAATAATGCTGCT 5325
QY 641 ggaacaaatcccaac 654
DB 5326 GGTTCATCTCTAC 5339

Search completed: June 23, 2000, 22:53:01
Job time: 40577 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:44:52 ; Search time 274.21 Seconds
(without alignments)

465,028 Million cell updates/sec

Title: US-09-180-798-24

Perfect score: 981
Sequence: 1 agtgcgaataattactgttg.....tcgcctaattactcaact 981

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5E_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/5F_COMB.seq:*
7: /cgn2_6/ptodata/2/ina/5G_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	6.8	3921	4 US-08-567-375-3	Sequence 3, Appl1
2	67	6.8	3921	4 US-08-587-680A-3	Sequence 3, Appl1
3	67	6.8	5992	3 US-08-475-891A-3	Sequence 3, Appl1
4	60.2	6.1	2075	1 US-08-238-163-3	Sequence 3, Appl1
5	60.2	6.1	6256	3 US-08-475-891A-1	Sequence 1, Appl1
6	60.2	6.1	6256	4 US-08-567-375-1	Sequence 1, Appl1
7	60.2	6.1	6256	4 US-08-587-680A-1	Sequence 1, Appl1
8	56.8	5.8	1058	1 US-08-238-163-1	Sequence 1, Appl1
9	56.8	5.8	5733	3 US-08-473-553A-1	Sequence 1, Appl1
10	54.2	5.5	1554	4 US-08-587-680A-24	Sequence 24, Appl1
11	48.2	5.0	831	4 US-08-367-375-15	Sequence 15, Appl1
12	48.4	4.9	3905	4 US-08-666-271-1	Sequence 4, Appl1
13	47.4	4.8	7218	1 US-08-232-463-14	Sequence 14, Appl1
14	47.4	4.8	7218	1 US-08-244-646-14	Sequence 14, Appl1
15	41.4	4.2	1116	2 US-08-592-936B-20	Sequence 20, Appl1
16	41.4	4.2	2917	4 US-08-111-573-20	Sequence 20, Appl1
17	41.4	4.2	2917	4 US-08-244-646-16	Sequence 16, Appl1
18	40.8	4.2	792	2 US-08-189-256A-16	Sequence 16, Appl1
19	38.4	3.7	1448	3 US-08-602-264A-11	Sequence 11, Appl1
20	36.4	3.5	2088	3 US-08-724-394A-20	Sequence 20, Appl1
21	34.6	3.5	246240	3 US-08-724-394A-20	Sequence 20, Appl1
22	34.6	3.5	246240	3 US-08-724-394A-21	Sequence 21, Appl1
23	34.6	3.5	246240	3 US-08-724-394A-22	Sequence 22, Appl1
24	34.6	3.5	246240	3 US-08-623-906A-16	Sequence 16, Appl1
25	34	3.4	251	3 US-08-508-786-2	Sequence 2, Appl1
26	33.8	3.4	1988	6 PCT-US96-12158-2	Sequence 2, Appl1
27	33.8	3.4	1988	6 PCT-US96-12158-2	Sequence 2, Appl1

C	28	33.8	3.4	2010	3	US-08-508-786-1	Sequence 1, Appl1
C	29	33.8	3.4	2010	6	PCT-US96-12158-1	Sequence 1, Appl1
C	30	33.6	3.4	2363	1	US-08-096-947-2	Sequence 2, Appl1
C	31	33.6	3.4	2363	2	US-07-919-140B-2	Sequence 2, Appl1
C	32	33.6	3.4	2363	6	PCT-US93-06939-2	Sequence 2, Appl1
C	33	33.3	3.4	291	2	US-08-442-063A-32	Sequence 32, Appl1
C	34	33.3	3.4	426	2	US-08-442-063A-35	Sequence 35, Appl1
C	35	33.3	3.4	567	2	US-08-442-063A-38	Sequence 38, Appl1
C	36	33.3	3.4	711	2	US-08-442-063A-41	Sequence 41, Appl1
C	37	33.3	3.4	849	2	US-08-442-063A-44	Sequence 44, Appl1
C	38	33.3	3.4	924	2	US-08-442-063A-47	Sequence 47, Appl1
C	39	33.3	3.4	1002	2	US-08-442-063A-26	Sequence 26, Appl1
C	40	33.3	3.4	1026	1	US-08-272-919-1	Sequence 1, Appl1
C	41	33.3	3.4	1026	2	US-08-619-916-1	Sequence 1, Appl1
C	42	33.3	3.4	1026	6	PCT-US95-08342-1	Sequence 1, Appl1
C	43	33.3	3.4	1593	7	5340934-3	Sequence 1, Appl1
C	44	33.3	3.4	4673	1	US-07-638-431-1	Sequence 1, Appl1
C	45	33.3	3.4	4673	6	PCT-US92-00018-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-567-375-3
: Sequence 3, Appl1 US/08567375
: Patent No. 5952485
: GENERAL INFORMATION:
: APPLICANT: Ronald, Pamela C.
: APPLICANT: Wang, Guo-Liang
: APPLICANT: Song, Wen-Yuang
: APPLICANT: Szabo, Veronique
: TITLE OF INVENTION: Procedures and Materials for Confering
: TITLE OF INVENTION: Disease Resistance in Plants
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/567,375
: FILING DATE: 04-DEC-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/004,645
: FILING DATE: 29-SEP-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/475,891
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/373,375
: FILING DATE: 17-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34,774
: REFERENCE/DOCKET NUMBER: 023070-058930
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3921 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product= "Xa-21"
US-08-567-375-3

Query Match 6.8%; Score 67; DB 4; Length 3921;
Best Local Similarity 49.9%; Pred. No. 4.7e-10;
Matches 169; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 344 aactctctgacatcttgccctgagcttgaggagcttgaaacttaacatctgacag 403
DB 1159 AAGATCAAGAGACATATCCAGAGATATTTGGCAATCTTTTGGCTTCAACAATCTCTAT 1218
QY 404 ctctacaacaaacacatcccaaggaactacatcttcgaacttggaaacttgaagatctc 463
DB 1219 CTCTGCAACACAAATTTTCAGAGGGTCTCTCCATCATCGTTGGCAGGCTTAAAACTTA 1278
QY 464 atcagcttgatctgtacaacaaacatcttaagagtagtcccaactcttgggaa 523
DB 1279 GGCATTTCTACTCGCTACGAAACACTTGAAGCGTTGATCCCGTTGGCCATAGGAAT 1338
QY 524 ttgaagctctgctcttcttaacgcttaatgacacccgattgacgggtccatccctaga 583
DB 1339 CTTACTGAACTTAATATCTTACTGCTCGGACCAACAATTCAGTTGGATACATAC 1398
QY 584 gcactacgagcaatcccaagccttaagttgtgaagctcgaagaatgattgtgga 643
DB 1399 ACACCTCAAACTCACAACAACTTGTGTGATAGGCTTCAACTAATACCTTAGTGT 1458
QY 644 acaatcccaacaaacgagcccttgcacattcctta 682
DB 1459 CCAATACCCAGTGAATTAATCAATTTCAACACTATCA 1497

RESULT 2

US-08-587-680A-3
Sequence 3, Application US/08587680A
Patent No. 587434
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Conferring
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,680A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645

FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/567,375
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058940US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product= "Xa-21"
US-08-587-680A-3

Query Match 6.8%; Score 67; DB 4; Length 3921;
Best Local Similarity 49.9%; Pred. No. 4.7e-10;
Matches 169; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 344 aactctctgacatcttgccctgagcttgaggagcttgaaacttaacatctgacag 403
DB 1159 AAGATCAAGAGACATATCCAGAGATATTTGGCAATCTTTTGGCTTCAACAATCTCTAT 1218
QY 404 ctctacaacaaacacatcccaaggaactacatcttcgaacttggaaacttgaagatctc 463
DB 1219 CTCTGCAACACAAATTTTCAGAGGGTCTCTCCATCATCGTTGGCAGGCTTAAAACTTA 1278
QY 464 atcagcttgatctgtacaacaaacatcttaagagtagtcccaactcttgggaa 523
DB 1279 GGCATTTCTACTCGCTACGAAACACTTGAAGCGTTGATCCCGTTGGCCATAGGAAT 1338
QY 524 ttgaagctctgctcttcttaacgcttaatgacacccgattgacgggtccatccctaga 583
DB 1339 CTTACTGAACTTAATATCTTACTGCTCGGACCAACAATTCAGTTGGATACATAC 1398
QY 584 gcactacgagcaatcccaagccttaagttgtgaagctcgaagaatgattgtgga 643
DB 1399 ACACCTCAAACTCACAACAACTTGTGTGATAGGCTTCAACTAATACCTTAGTGT 1458
QY 644 acaatcccaacaaacgagcccttgcacattcctta 682
DB 1459 CCAATACCCAGTGAATTAATCAATTTCAACACTATCA 1497

RESULT 3

US-08-475-891A-3
Sequence 3, Application US/08475891A
Patent No. 5859339
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Conferring
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: us/08/475,891A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 02370-05891005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5992 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(512..3149, 3993..4393)
OTHER INFORMATION: /product= "RRK-B"
OTHER INFORMATION: /note= "Xa21 xanthomonas spp. disease
OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
OTHER INFORMATION: sativa)"
US-08-475-891A-3
```

```

Query Match          6.88; Score 67; DB 3; Length 5992;
Best Local Similarity 49.9%; Pred. No. 5.8e-10;
Matches 169; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
```

```

QY 344 aaccctctggacatctcgccgcctgagcttggaagcttgaaacattacagatctcagag 403
   || || || || || || || || || || || || || || || || || || || || ||
DB 1670 AAGATCAGAGGAGGATTCGAGAGATTTGGCAATTTATTGGCTTACAACTCTCTAT 1729
QY 404 ctctcaaaaaaacatccaaaggaactatccttcggaacttggaacttcgaagactc 463
   || || || || || || || || || || || || || || || || || || || || ||
DB 1730 CTCTTCACAACAATTTCAGAGGGTCTCTTCATCATCGTTGGCAGGCTTAAACTTA 1789
QY 464 atcagcttgatctgttaacaaacatcttcagagatagttcccaattcttgggaaaa 523
   || || || || || || || || || || || || || || || || || || || || ||
DB 1790 GGCACTTCTACGCGCTACGAAACAACTTGAGCGGTTCGATCCCGTTGGCCATAGGAAT 1849
QY 524 ttgaagctctcgctcttcttaagccttaataacacagcattgacgggtccatccctaga 583
   || || || || || || || || || || || || || || || || || || || || ||
DB 1850 CTACTGAGACTTAATATTTACTGCTCGGACCAACAATTCACTGATGGATACCATAC 1909
QY 584 gcactcaagcgaatcccaagccttaagttgttgaagcttcaagaatgattgtgtgga 643
   || || || || || || || || || || || || || || || || || || || || ||
DB 1910 AACTCTCAAACTCACAACCTTGTTGTCATAGGCTTTCACTAATCACTTAGTGTG 1969
QY 644 acaatcccaaaaacagcgccttgcctcacattcctta 682
   || || || || || || || || || || || || || || || || || || || || ||
DB 1970 CCAATACCCAGTGAATTTTCAATATTCAAACATATCA 2008
```

```

RESULT 4
US-08-238-163-3
; Sequence 3, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAYTICH, John M.
; APPLICANT: POWELL, Ann
; APPLICANT: STOLT, Henrik
```

```

TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,163
FILING DATE: 03-MAY-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 2307E-540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 421..1401
US-08-238-163-3
```

```

Query Match          6.1%; Score 60.2; DB 1; Length 2075;
Best Local Similarity 49.0%; Pred. No. 3.3e-08;
Matches 221; Conservative 0; Mismatches 223; Indels 7; Gaps 2;
```

```

QY 204 tctacgctcttcgcgcgagtttgacagatccagacatgtctccagagcttggatccaa 263
   || || || || || || || || || || || || || || || || || || || || ||
DB 513 TCTCAAAATAAAGAAAGACTTAGCAATCTTACCAATTTAGCTTCAATGGGATCCAAAGAC 572
QY 264 ctcttgtaaatcccttgtaacctgttcaatgtcaacctgttaaccaagaacacgctcacc 323
   || || || || || || || || || || || || || || || || || || || || ||
DB 573 AGATTGCTGTACGCTACGTCATCAATAATGTGACCGGAAC--CAACCGGATTAATG 628
QY 324 gtgtgatttgggaattccaacctctctgacatcttggccttgagcttgggaagcttg 383
   || || || || || || || || || || || || || || || || || || || || ||
DB 629 CTCTACCGCTTCTCCAGGCCAATATCTCCGCCAANTCCGGACCGCTTCGAGAGACTTC 688
QY 384 aacattacagtatctcagcctcaca--aaacaacatccaaagaaactataaccttcg 440
   || || || || || || || || || || || || || || || || || || || || ||
DB 689 CATATCTGGAACATTTGGAATTTTCATCATGTACTATCTCACCGGAACAAATCTCACCTG 748
QY 441 aacttggaaactctgaagaatctcagcgtgtgatactgtacaaacaacatcttaacagga 500
   || || || || || || || || || || || || || || || || || || || || ||
DB 749 CAATTGCGAAGCTCACAATCTCAAAATGTTAAGCTCAAGCTTCACTTAACCTTACAGTC 808
QY 501 tagttcccaactcttgggaattgaagctctcgtcttctttaaagcgttaatgaacacc 560
   || || || || || || || || || || || || || || || || || || || || ||
DB 809 CGATCCCTGAATTCCTTAGTACGCTGAAGAAATTGAGCTGCTCGATGGAATTAACATC 868
QY 561 gatgacggtcacaatccctcagcagcctcagcgaatcccaacgcttaagttgttgaag 620
   || || || || || || || || || || || || || || || || || || || || ||
DB 869 AATTACCGGAACAATCTCTTCTCTCTCTCAAGCTTCGGAATTTGCTAGCATGACT 928
QY 621 tctcaagcaatgatttgttggaaacatccc 651
   || || || || || || || || || || || || || || || || || || || || ||
```

```

Db      929 TAGATCGTACAAACTCACGGAGCAACAATACC 959
RESULT          5
US-08-475-891A-1
Sequence 1, Application US/08475891A
Patent No. 5859339
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
TITLE OF INVENTION: Procedures and Materials for Confering
NUMBER OF INVENTIONS: Disease Resistance in Plants
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,891A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 02370-058910US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1648..4383, 5178..5513)
OTHER INFORMATION: /product="RRK-F"
OTHER INFORMATION: /note="xaz1 xanthomonas spp. disease
OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
OTHER INFORMATION: sativa)"
US-08-475-891A-1
Query Match          6.1%; Score 60.2; DB 3; Length 6256;
Best Local Similarity 51.1%; Pred.No.5,6e-08;
Matches 168; Conservative 0; Mismatches 158; Indels 3; Gaps 1;
QY      327 tggatttggaaatccaacctctcgagcatcttcgctgtgagtgggaactgtgaac 386
         ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      2735 ttgcaccttggaagaataatmaacttgggggaggtttctctaatttcgattcttccaa 2794
QY      387 att---taacgatcttagagctctacaacaacaacatccaaggactataacttcgaaac 443
         || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2795 cttcgcttagttttcttcgcacttgatttgatatagatcacacgaacattccaaagcata 2854
QY      444 ttgaataatcgaaagaaatctatcagcttggatctgtgaacaanaaatcttaagggatag 503
         ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      2855 ttggcaatcttatgtgccttacaaactctatcttatcttgcaaacacaaatttcagaggctcac 2914

```

```

OY 504 ttccactctttaggaaatagatgagctctgtctctttagagcttaatgacaacgat 563
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2915 ttccatcattcgtttggacggctttgaaatttagccttttagtcccttaccaaaacact 297
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 564 tgacgcgtccatcccttagagcactcagcgaatcccaagccttaagtgttgacgtct 623
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2975 tgagcggttgcattcccatgtgacccataggaatcttactgaacttaattcttactgctcg 303
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 624 caagcagatgattgtgtggaacaatccca 652
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3035 gcaccmaacmaattcagttggttgatacca 3063
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-08-567-375-1
; Sequence 1, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice
; OTHER INFORMATION: (Oryza sativa)"
US-08-567-375-1

```


SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,163
FILING DATE: 03-MAY-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 2307E-540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1058 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1023
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1058
OTHER INFORMATION: /standard_name="Pear FGP cDNA"
US-08-238-163-1

Query Match 5.8%; Score 56.8; DB 1; Length 1058;
Best Local Similarity 52.0%; Pred. No. 2.3e-07;
Matches 127; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 342 caaacctctgagcaccctggcgctgagcttggaagcttggaattcactctag 401
DB 347 CCAATCTCACTGGCCCAATCCACCCGCAATGCGCAAGCTCAAGACTCTCTCA 406
QY 402 agctctcaaaaacaacatccaagaactatacctccgaacttggaattcgaagaatc 461
DB 407 GGCCTCAGCTGACCAACOTCTCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 466
QY 462 tcatcaagcttgatctgtacacaacaatttcaagaggaagtccactctcttgga 521
DB 467 TCAATCTCTGACCTCTCTCTCAACAACCTCACCGGTGCGCATCCCAAGCTCTTCTG 526
QY 522 aattgaagctctgtctcttttaagcgttaatgaacaacgattgaccggtccactc 581
DB 527 AGCTCCCAAACTGGGGCTCTTCTGCTAGACCGCAATAGCTCACAGCTCATATTCCGA 586
QY 582 gaagc 585
DB 587 TATC 590

RESULT 9
US-08-473-553A-1
Sequence 1, Application US/08473553A
Patent No. 5859338
GENERAL INFORMATION:
APPLICANT: Meyerowitz, Elliot M.
APPLICANT: Clark, Steven E.
APPLICANT: Williams, Robert W.
TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,553A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5733 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 2434..5037
FEATURE:
NAME/KEY: CDS
LOCATION: 5117..5467
US-08-473-553A-1

Query Match 5.6%; Score 54.6; DB 3; Length 5733;
Best Local Similarity 51.0%; Pred. No. 2.3e-06;
Matches 129; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 406 ctacaanaacaacatccaagaactatacctccgaacttggaattcgaagaatctcat 465
DB 3120 CTACTACACACAGCTACACCGGTGCTGTCACGCGAGTTCGGTTTACAAGCTTGA 3179
QY 466 cagcttgatctgtacacaacaactcttaagagtagttccactctcttggaatt 525
DB 3180 GATCTTCGACATGGGAGCTTACACTCACCGGAGAGATTCCGACGAGTTAAGTAACT 3239
QY 326 gaagctctgtctcttttaagcgttaatgaacaacgattgaccggtccactagac 585
DB 3240 GAAACATCTACACTCTGTTCTTACATCAACAACTTAACCGGTATACACCGGA 3299
QY 586 actcagcgaatcccaagccttaagtggtgacgtctcaagaagtattggtgaac 645
DB 3300 GCTTCCGGTTAGTCACTGAATCTCTCGATTATCATCATCATCAATCAACCGGAGA 3359
QY 646 aatcccaacaac 658
DB 3360 AATCCCTCAAGC 3372

RESULT 10
US-08-587-680A-24
Sequence 24, Application US/08587680A
Patent No. 5977434
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco


```

US-08-666-271-4
; Sequence 4, Application US/08666271
; Patent No. 5920000
;
GENERAL INFORMATION:
;
APPLICANT: JONES, JOINTHAN D
APPLICANT: HAMMOND-KOSACK, KIM E
APPLICANT: THOMAS, COLWYN M
APPLICANT: JONES, DAVID A
TITLE OF INVENTION: PLANT PATHOGEN RESISTANCE GENES AND USES
TITLE OF INVENTION: THEOREF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
;
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/666,271
FILING DATE: 19-SEP-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/02812
FILING DATE: 23-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326428.1
FILING DATE: 24-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9409363.0
FILING DATE: 11-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2880 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-666-271-4

```

	Query Match	Similarity	4.9%	Score 48.4	DB 4	Length 2880
	Best Local	Similarity	48.5%	Pred. No. 0.0001		
	Matches	133	Conservative	0	Mismatches 14	Indels 0
					Gaps 0	
Qy	397	tctagagctctcaaaaacacatccaagaactacccctcgaacttgaatctgaa	456			
Db	2073	TATCATCTCTCTCAAGAACAGATTGAAAGCATATTCCAAACATATTGAGCATCTGT	2132			
Qy	457	gaatcatcagcttggatctgttaacaaacaatcttaaggatagttcccaactctt	516			
Db	2133	TGGACTTCGTACGTTGAACTTCTTCACAAATGCTTGGAAAGTCATATACCGCATATT	2192			
Qy	517	gggaaaatgaaagctctgtgctcttttaagcgtttaatgaaacagattgaccgltcat	576			
Db	2193	TCAAAATTATACGATCTCGAATCTTTGATCTCTCATCTAATAAATACGGGAGAAAT	2252			
Qy	577	cctctgagagactccaaggcaatcccaagccttaaatgttgtgacggtctcaagaatgat	636			
Db	2253	TCCGACAGACGCTTGATCCCTCCTACATCTCTTGAAGCTTAAATCTCTTCACATATATCT	2312			

Qy 637 gtgtggaacaatcccaacaaacgacccttgc 670
||| ||| | | ||| |
Db 2313 TGTGGATGCATCCCCAAAGAAACAATTTGAT 2346

```

RESULT 13
US-08-666-271-1
Sequence 1, Application us/08666271
Patent No. 5920000
GENERAL INFORMATION:
APPLICANT: JONES, JONATHAN D
APPLICANT: HAMMOND-KOSACK, KIM E
APPLICANT: THOMAS, COLWYN M
APPLICANT: JONES, DAVID A
TITLE OF INVENTION: PLANT PATHOGEN RESISTANCE GENES AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/666,271
FILING DATE: 19-SEP-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/02812
FILING DATE: 23-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326428.1
FILING DATE: 24-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9409363.0
FILING DATE: 11-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3905 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 898..3489
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 967..3486
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 898..966
US-08-666-271-1

```

Query Match	4.99;	Score 48.4;	DB 4;	Length 3905;
Best Local Similarity	48.58;	Pred. No. 0.00012;		
Matches 133;	Conservative	0;	Mismatches 141;	Indels 0;
				Gaps

397 tttagagctctcaaaaacatccaaaggaactatcccttcgcgaacttgyaaatctgaa 456

Db 2913 TATCAATCTCTCAAGACAGATTGTGAGCTATTTCCAGCATTAATTGGAGATCTTGT 2972
QY 457 gaatcatcaagcttgatctgtacaacaacactacaggaatgttccactcttt 516
Db 2973 TGGACTTCGTACGTTGAATCTTGCTCACATGTCTTGAAGGTCAATATACCGCATCATCT 3032
QY 517 gggaaattgaagctctgtcttttaagcgttaataagacacgaattgacggctcaat 576
Db 3033 TCAAAATTTATTCAGTACTTCGAATCTTGGATCTCTCATCTAATAAATCAACGAGAAAT 3092
QY 577 ccctagagcaactcagcgcaactccaagccttaagattgttgacgtctcaagcaatgatt 636
Db 3093 TCCGACAGAGTTGCATCCCTCAATCTCTGAAGTCTTAAATCTCTCAACAATCATCT 3152
QY 637 gtgtggaacaatcccaacaacggaaccttgcct 670
Db 3153 TGTGTGATGATCCCAAGAAACAAATTTGAT 3186

RESULT 14

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHREIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9p-Fls
; US-08-232-463-14

Query Match 4.8%; Score 47.4; DB 1; Length 7218;
Best Local Similarity 3.5%; Pred. No. 0.00032;
Matches 12; Conservative 194; Mismatches 135; Indels 0; Gaps 0;

QY 12 tttagttgtcttctctctcttcttgaagaatttccttactctcaattcctttc 71
Db 1115 YY 1174
QY 72 gattccctcttaacctcgaagctacatgagctcgaactatcggtggagc 131
Db 1175 YY 1234
QY 132 tcttcgaagcttgtaaccttaaccttagcttgattcaactggtcgaagaactcg 191
Db 1235 YY 1294
QY 192 aagagatgctctcctacgctcttcgcccgaagttgacaatcagacactgtccca 251
Db 1295 YY 1354
QY 252 gctggatccaactctgttaactctgtactgctcatgcatgtaacctgaaccaaga 311
Db 1355 YY 1414
QY 312 accgagctcaactgctgtgagatttggaattcaacctctct 352
Db 1415 YYYYYYYYYYYYYYYYYYGTACCAATCTCTATCTCT 1455

RESULT 15

US-08-244-646-14
; Sequence 14, Application US/08244646
; Patent No. 5744692
; GENERAL INFORMATION:
; APPLICANT: Cervone, Felice
; APPLICANT: De Lorenzo, Giulia
; APPLICANT: Salvi, Giovanni
; APPLICANT: Albersheim, Peter
; APPLICANT: Darvill, Alan
; APPLICANT: Bergmann, Carl
; TITLE OF INVENTION: Nucleotide Sequences Coding An
; TITLE OF INVENTION: Endopolygalacturonase Inhibitor
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sally A. Sullivan
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,646
; FILING DATE: 06-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM 91A 000915
; FILING DATE: 06-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/IT/00158
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sullivan, Sally A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 19-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)498-8080
; TELEFAX: (303)498-8089
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1116 base pairs
; TYPE: nucleic acid

```
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Phaseolus vulgaris
STRAIN: Saxa
IMMEDIATE SOURCE:
CLONE: lambda PGIP-3.3
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1026
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1027..1116
US-08-244-646-14
```

```
Query Match 4.28; Score 41.4; DB 2; Length 1116;
Best Local Similarity 48.18; Pred. No. 0.0072;
Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 409 caaaacacatccaaggaactatatacttcgaacttgaatctgaagaatctcatcag 468
DB 342 CATCAATAACCTGTCGGTCCAAATCCCCCGCCATCGCTAACTCACCACCACTCCACTA 401
QY 469 ctggatctgtacaacaacaatctacagggatgtcccaactcttgggaaattgaa 528
DB 402 TCTTATATACACTACACCAATGTCTCGGCGCAATACCCGATTTCTGTCAAGATCAA 461
QY 529 gctctgtgcttttaagsgcttaatgaacacgatgaacgggtgcgaatccctagagcaat 588
DB 462 AACCCCTGTCACCTTGACTTCTCTACACGCCCTCTCCGCACTCCCTCCCTCCAT 521
QY 589 cagggcaatcccaagccttaaaagtgtgaagctctcaagcaatgatttgttgaaacaat 648
DB 522 CTCTTCTCTCCCAACCTCGAGGATCATTCGACGCAACGAATCTCCGGCCCAT 581
QY 649 ccc 651
DB 582 CCC 584
```

Search completed: June 23, 2000, 22:45:22
Job time: 40268 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:32 ; Search time 6198.48 Seconds
(without alignments)
641.482 Million cell updates/sec

Title: US-09-180-798-24
Perfect score: 981
Sequence: 1 agtgtgagtaattagtgtg.....tcgcctaatactaccact 981

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202661650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query

SUMMARIES

NO.	Score	Match	Length	DB	ID	Description
1	403	41.1	469	63	A1992868	A1992868 701493826
2	384.2	39.2	556	23	R89998	R89998 16353 Lambd
3	359.6	36.7	534	33	AA394359	AA394359 25942 Lam
4	353.2	36.0	447	23	H36800	H36800 14929 Lambd
5	343.6	35.0	401	37	AA712221	AA712221 31949 Lam
6	335.8	34.2	353	23	H37300	H37300 15429 Lambd
7	333	33.9	364	42	A1100683	A1100683 33804 Lam
8	328.8	33.5	336	20	Z34187	Z34187 ATIS3221 Ve
9	315.6	32.2	348	42	A1100682	A1100682 33803 Lam
10	312.6	31.9	608	81	AW443205	AW443205 EST308135
11	311.6	31.8	720	64	AW031110	AW031110 EST274417
12	305	31.1	337	20	Z34606	Z34606 ATFS3366 Ve
13	298.2	30.4	339	42	A1100679	A1100679 33800 Lam
14	293.8	29.9	646	74	AW221939	AW221939 EST296750
15	292.8	29.8	657	80	AW350720	AW350720 GM210009A
16	291.8	29.7	336	23	H37296	H37296 15425 Lambd
17	288.8	29.4	511	74	AW221278	AW221278 EST297747
18	286.2	29.2	714	64	AW030188	AW030188 EST273443
19	276.8	28.2	334	23	H37195	H37195 15324 Lambd
20	272.8	27.8	695	64	AW038168	AW038168 EST279825
21	270.4	27.6	465	45	A1352795	A1352795 MB61-10D
22	268	27.3	599	63	AW011134	AW011134 S117B03 P
23	267.4	27.3	599	74	AW219797	AW219797 EST302279
24	267.2	27.2	305	25	N65416	N65416 20456 Lambd
25	260.6	26.5	637	80	AW350549	AW350549 GM210009A
26	256	26.1	336	42	A1100680	A1100680 33801 Lam
27	255.8	26.1	676	74	AW220075	AW220075 EST302558
28	254	25.9	442	42	A1100481	A1100481 34856 Lam
29	252.8	25.8	430	79	AW307218	AW307218 S154C07 Y
30	251.8	25.7	569	59	A1775448	A1775448 EST256548
31	244	24.9	522	46	A1441759	A1441759 S882D08 Y
32	241.6	24.6	360	42	A1100685	A1100685 33806 Lam
33	241.2	24.6	545	64	AW037836	AW037836 EST279465
34	239.4	24.4	667	79	AW185847	AW185847 S606D04 Y
35	231.4	23.6	515	79	AW279515	AW279515 S190E09 Y
36	231.2	23.6	612	47	A1496325	A1496325 S805C09 Y
37	230.4	23.5	541	64	AW036865	AW036865 614019G10
38	229.6	23.4	619	64	AW040482	AW040482 EST283442
39	227.2	23.2	304	36	T21150	T21150 3158 Lambd
40	224.2	22.9	568	47	A1487272	A1487272 EST245594
41	223.6	22.8	423	35	C22371	C22371 C22371 R1CE
42	221.8	22.6	570	50	A1676939	A1676939 605047A07
43	218.2	22.2	616	42	AF074734	AF074734
44	215.8	22.0	540	64	AW033599	AW033599 EST277170
45	214.6	21.9	258	42	A1100678	A1100678 33799 Lam

ALIGNMENTS

RESULT	1	LOCUS	469 bp	EST	08-SEP-1999
DEFINITION	A1992868	701493826 A. thaliana	Ohio State clone set	Arabidopsis thaliana	
ACCESSION	A1992868	CDNA clone 701493826, mRNA sequence.			
VERSION	A1992868.1	GI:5839773			
KEYWORDS	EST				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1	(bases 1 to 469)			
AUTHORS	Chen, J., Komiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mounoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Rastbury, K., Borrell, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrta, A., Murry, L., Turner, C., Kikorian, S., Elder, L., and				

TITLE
JOURNAL
COMMENT

Hanson, D.
Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)
On Dec 20, 1995 this sequence version replaced gi:1135328.
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES

source

1..469
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="701493826"
/clone_1ib="A. thaliana, Ohio State clone set"
/note="CDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."

BASE COUNT 120 a 124 c 89 g 136 t

Query Match 41.1%; Score 403; DB 63; Length 469;
Best Local Similarity 96.5%; Pred. No. 8,6e-92;
Matches 412; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY	57	ctcaatctcttctgattccctccttaaacctcgaagctcagcgctcgaa	116
DB	43	CTCAAAATTCCTATGATTAATCTCTCTTGGACCTCGAAGCTCAAGGCTTCGAA	102
OY	117	actacgttggaagctcttcagcttcgttaacaccttaacaccttagcttgatcaactg	176
DB	103	ACTATGCTGGAGCTCTTGGACCTTCGTAATCTTAACCTAGCTTGAATCACTGG	162
OY	177	tgaagcaaacctccgaagagatgctctcagctcttcgacggaattggaagatccag	236
DB	163	TGGAAGCAACTCCGAAGAGATGCTTTACGCTTCCCGAGTTTAAACAGATCCG	222
OY	237	accatgctctccagaagcttggaatcactctgttaacaccttgatcctggttcacatga	296
DB	223	ACCATGCTTCCAGAGCTGGAGATCCAACTTGTATCTTACTGTTCCATGCTCA	282
OY	297	ccttgtaaccaagaacacgcgtcactcgttggaatttggaatctaaacacctctgac	356
DB	283	CCTGTAACCAAGACACCGGCTCACTGCTGATTTGGGAATTCGAACCTCTCGAC	342
OY	357	atcttgagctgagcttggaagcttgaacattacagttactaaggtctataaanaa	416
DB	343	ATCTTGCGCTGAGCTTGGAGAGCTTGAACATTACAGTATCTAGAGCTTACATTA	402
OY	417	acatccaagaacataccttcgaacttggaatctgaagaatctcaatcagcttgatc	476
DB	403	ACATCCAAGAACATATACCTTCCGAATCTGGAATCTGAAGATCTCATGAGCTTGATC	462
OY	477	tgtaaca 483	
DB	463	TGTACAA 469	

RESULT	2	LOCUS	556 bp	EST	30-DEC-1997
DEFINITION	R89998	16353 LambdA-PR12 Arabidopsis thaliana	CDNA clone 187H577, mRNA		
ACCESSION	R89998	sequence.			
VERSION	R89998.1	GI:957538			
KEYWORDS	EST				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1	(bases 1 to 556)			
AUTHORS	Chen, J., Komiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mounoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Rastbury, K., Borrell, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrta, A., Murry, L., Turner, C., Kikorian, S., Elder, L., and				

REFERENCE 1 (bases 1 to 556)
 AUTHORS Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.
 TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 COMMENT On Apr 14, 1993 this sequence version replaced gi:693023.
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@lhm.cl.msu.edu
 Seq primer: T7 dye primer.

FEATURES

source
 1..556
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="187517"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."
 BASE COUNT 151 a 118 c 123 g 146 t 18 others
 ORIGIN

Query Match 39.2%; Score 384.2; DB 23; Length 556;
 Best Local Similarity 92.9%; Pred. No. 4.9e-87;
 Matches 434; Conservative 0; Mismatches 27; Indels 6; Gaps 3;
 QY 347 cctcttgacatcttgccctgagctgagctggaacattgaacatctagaagctc 406
 Db 1 CTCTGTGACATCTTGGCCGAGCTTGAGAGCTTGAACATTTACAGTATCTAGAGCTC 60
 QY 407 tacaaaaaacatccaagaactatccctccgaacttggaacttgagaatccctc 466
 Db 61 TACAAAAAACATCCACGAGCACTATACCTCCGAACCTTGAATCTGAAGATCTCATC 120
 QY 467 agcttgatctgtacaacaacatctacagggatagttcccaactcctcttggaatc 526
 Db 121 AGCTTGATCTGTACACCAACATCTTACAGGGATAGTCCACTCTTGGAAAAATTG 180
 QY 537 aagctctggtcttttaaggcttaatgacaacggatgacggtccaatccctagaag 586
 Db 181 AAGCTCTGGCTTTTACGGCTTAATGACACCGATGACGGGGCCAAATCCTTACAGCA 240
 QY 587 ctccagcaatcccaagccttaaggttgtagcgtccagcaatgattgtgtgaaaca 646
 Db 241 CTCACGTCAATCCCAAGCCTTAAGGTGTGTGATGTCACAGCAATGATTGTGTGNA 300
 QY 647 atcccaacaaggaaccttgctacatccctttacagaacttgagaacaaccgaga 706
 Db 301 ATCCCAACAAAGGACCTTTTGTCTCACATTCCTTTACAGAACTTTGAGNACAACCCGAG 360
 QY 707 ttgg-agggacaggaattactcgtctgctgcaagctagcaactaacgtgcaaccga 765
 Db 361 TTGGAGGAGGACGGAATTACTCGGTCTTGCAGCTNCGCAATTAATGACGAC---TGA 417
 QY 766 ctggcaaaacctgaataatgaagaattggygggtgacctgtlaagaac 812

Db 418 ATTGAAAAACCTGAAA--TGAGATTGGGGGTGNCCTTTAAGGCAC 462
 RESULT 3
 AA394359
 LOCUS
 DEFINITION 25942 Lambda-PRL2 Arabidopsis thaliana cDNA clone 305G117 3', mRNA
 sequence.
 ACCESSION AA394359
 VERSION AA394359.1 GI:2047570
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 REFERENCE 1 (bases 1 to 534)
 AUTHORS Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.
 TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 COMMENT On May 18, 1995 this sequence version replaced gi:811121.
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@lhm.cl.msu.edu
 Seq primer: T7.
 Location/Qualifiers
 1..534
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="1305617"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

BASE COUNT 151 a 114 c 125 g 126 t 18 others
 ORIGIN

Query Match 36.7%; Score 359.6; DB 33; Length 534;
 Best Local Similarity 89.3%; Pred. No. 8.1e-81;
 Matches 424; Conservative 0; Mismatches 44; Indels 7; Gaps 4;
 QY 465 tcaactggatctgtacaacaacatcttaacggatagttcccaactcttgggaat 524
 Db 1 TCAGCTTGATCTGTACACCAACATCTTACAGGGATAGTCCACTTTTGGAAAAAT 60
 QY 525 tgaagctctggtcttttttaaggcttaatgacaacggatgacggtcccaatccctag 584
 Db 61 TGAAGCTCTGGCTTTTACGGCTTAATGACAAACGATTTGACCGGTCCAAATCCTTAG 120
 QY 585 cactcagggcaatcccaagccttaaggttgtagcgtcccaagcaatgattgtgtgaa 644
 Db 121 CACTCAGGCAATCCCAAGCCTTAAGGTGTGACGTCTCAAGCAATGATTTGTGGAA 180

Lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA.

BASE COUNT 103 a 105 c 78 g 99 t 16 others
ORIGIN

Query Match 35.0%; Score 343.6; DB 37; Length 401;
Best Local Similarity 92.1%; Pred. No. 8.8e-77;
Matches 363; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 83 cttaacctccgaagctacatgagctctcgaactatcgtggagctcttcagact 142
DB 1 ctttaaccnccgaagctacatgagctctcgaactatcgtggagctcttcagact 60
QY 143 tcttaaccctcaacctagcttagcttagcttagcttagcttagcttagcttagct 202
DB 61 tcttaaccctcaacctagcttagcttagcttagcttagcttagcttagcttagct 120
QY 203 cttaacgctcttcgagagcttagcttagcttagcttagcttagcttagcttagct 262
DB 121 cttaacgctcttcgagagcttagcttagcttagcttagcttagcttagcttagct 180
QY 263 actcttgtaactccttagcttagcttagcttagcttagcttagcttagcttagct 322
DB 181 actcttgtaactccttagcttagcttagcttagcttagcttagcttagcttagct 240
QY 323 cgttgagattgggaattcaaacctctcggagactcttcgagcttagcttagcttagct 382
DB 241 cgttgagattgggaattcaaacctctcggagactcttcgagcttagcttagcttagct 300
QY 383 gaacattacagatctagagctctacacacacacacacacacacacacacacacacac 442
DB 301 gaacattacagatctagagctctacacacacacacacacacacacacacacacacac 360
QY 443 ctgggaattcgaagatctcagcttagcttagcttagcttagcttagcttagcttagct 476
DB 361 ctgggaattcgaagatctcagcttagcttagcttagcttagcttagcttagcttagct 393

RESULT 6
H37300 353 bp mRNA EST 30-DEC-1997
LOCUS 15429 Lambda-PRU2 Arabidopsis thaliana cDNA clone 17901977, mRNA
DEFINITION
SEQUENCE
ACCESSION H37300
VERSION H37300.1 GI:906799
KEYWORDS EST
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 353)
REFERENCE
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On May 8, 1995 this sequence version replaced gi:801146.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.

Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313cn@dm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 353

FEATURES
source
/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone="1791977"
/note="Vector: lambda Zip-lox; Site 1: Sal; Site 2: Not; lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."
BASE COUNT 98 a 85 c 67 g 89 t 14 others
ORIGIN

Query Match 34.2%; Score 335.8; DB 23; Length 353;
Best Local Similarity 95.5%; Pred. No. 8.1e-75;
Matches 337; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 407 tacaaaaaacatccaaaggaactatactctcgaacttggaaatctgaagaatctcattc 466
DB 1 TACAAAAAACATCCAAAGGAACATACCTTCGAACTTGAATCTGAAGATCTCATC 60
QY 467 agcttgatctgtaacaaac 526
DB 61 agcttgatctgtaacaaac 120
QY 527 aagctctgctcttcttaacgactaagacacacacacacacacacacacacacacacac 586
DB 121 aagctctgctcttcttaacgactaagacacacacacacacacacacacacacacacac 180
QY 587 ctcaaggaatcccaagccttaagttgtgacgtctcaagaaatgattgtgtgagaca 646
DB 181 ctcaaggaatcccaagccttaagttgtgacgtctcaagaaatgattgtgtgagaca 240
QY 647 atcccaaacacagcagcccttgctcacatctcttcacacacacacacacacacacac 706
DB 241 ATCCCAACACAGCAGCCTTGTCTCACATCTCTTACAGACTTNGANACACCGGNGA 300
QY 707 ttggaaggaccggaattactcgtcttgcaagctacagacactaactgacactg 759
DB 301 ttggaaggaccggaattactcgtcttgcaagctacagacactaactgacactg 353

RESULT 7
A1100683 364 bp mRNA EST 21-AUG-1998
LOCUS 33804 Lambda-PRU2 Arabidopsis thaliana cDNA clone 1875XP 3', mRNA
DEFINITION
SEQUENCE
ACCESSION A1100683
VERSION A1100683.1 GI:3449393
KEYWORDS EST
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 364)
REFERENCE
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On May 8, 1995 this sequence version replaced gi:801146.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.

TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On May 5, 1995 this sequence version replaced gi:797711.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313c@msu.edu
The sequence entry for this EST has been reverse complemented and is being submitted in the sense orientation.
Seq primer: M13-21.

FEATURES
SOURCE

1. 364
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="187H5XP"
/note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using Oligo dT primed cDNA."
BASE COUNT 113 a 80 c 70 g 99 t 2 others
ORIGIN

Query Match

Best Local Similarity 96.4%; Score 333; DB 42; Length 364;
Matches 350; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 506 cccactctcttggaataatgaagctctgctcttttcaaggcttaataacacgagattg 565
DB 1 cccattcttttggaataatgaagctcttnt-gtctttttacggcttaataacacgagattg 59
QY 566 accggtccaatcccttagagcaatcagcgcaatcccaagccttaagttgttgagcttca 625
DB 60 acggggcccaatcmtatgacactcactgcacacccaagccttaagttgttgatctca 119
QY 626 agcaatgattgttggaacaatcccaacaacgagccttgcctcacatcctttacag 685
DB 120 agcaatgattgttggaacaatcccaacaacgagccttgcctcacatcctttacag 179
QY 686 aactttgagaacaacccgagattggaaggacggaattactcgtctctgcaagctagac 745
DB 180 aactttgagaacaacccgagattggaaggacggaattactcgtctctgcaagctagac 239
QY 746 aactaactgacccggaacacacgcaaacctgaaataaattgggggtgaccttg 805
DB 240 actaactgacccggaacacacgcaaacctgaaataaattgggggtgaccttg 299
QY 806 taagaacacttcacacattatcaaatatcacatctatctatgtaataatataatctg 865
DB 300 taagaacacttcacacattatcaaatatcacatctatctatgtaataatataatctg 359
QY 866 agt 868
DB 360 agt 362

RESULT 8
Z34187/c 234187 336 bp mRNA EST 14-JUN-1994
LOCUS

DEFINITION ATTS3221 Versailles-VB Arabidopsis thaliana cDNA clone VBVDH02.
ACCESSION Z34187
VERSION Z34187.1 GI:498544
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 336)
CNRS.
AUTHORS The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
CONTACT: Desprez T., Anselme J., Chiapello H., Rouze P., Caboche M., Hoite H.
COMMENT INRA Versailles
Laboratoire de Biologie Cellulaire
Route de Saint-Cyr 78026 Versailles Cedex, France
Email: therry@versailles.inra.fr.

FEATURES
SOURCE

1. 336
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia"
/db_xref="taxon:3702"
/clone="VBVDH02"
/clone_1lb="Versailles-VB"
/tissue_type="Whole seedlings"
/dev_stage="in vitro-grown etiolated seedlings, 5 days old"
/note="Vector: pBluescript"
BASE COUNT 89 a 62 c 65 g 119 t 1 others
ORIGIN

Query Match

Best Local Similarity 99.1%; Score 328.8; DB 20; Length 336;
Matches 330; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 633 attgtgtggaacaatcccaacaacgagccttctacattccttaacgaagattg 692
DB 333 attgtgtggaacaatcccaacaacgagccttctacattccttaacgaagattg 274
QY 693 agaacaacccgagattggaaggacggaattactcgtctcgaagctacgacactaact 752
DB 273 agaaacacccgagattggaaggacggaattactcgtctcgaagctacgacactaact 214
QY 753 gcaactgaaacacctggcaaacctgaataatgggggtgacctgtgaagac 812
DB 213 gcaactgaaacacctggcaaacctgaataatgggggtgacctgtgaagac 154
QY 813 acttcaccactttataaataatcacatctatataatgaatgaatataatgaatga 872
DB 153 acttcaccactttataaataatcacatctatataatgaatgaatataatgaatga 94
QY 873 acaaaaaaataagaagatcggaatcggttaataatcattgctcaattggaactcgagg 932
DB 93 acaaaaaaataagaagatcggaatcggttaataatcattgctcaattggaactcgagg 34
QY 933 tctgtatgtaaaatttctaattcgatttcgc 965
DB 33 tctgtatgtaaaatttctaattcgatttcgc 1

RESULT 9
A1100682 348 bp mRNA EST 21-AUG-1998
LOCUS A1100682
DEFINITION 33803 Lambda-PRL2 Arabidopsis thaliana cDNA clone 185L1XP 3', mRNA
sequence.
ACCESSION A1100682
VERSION A1100682.1 GI:3449392
KEYWORDS EST.
SOURCE thale cress.

ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsids.
1 (bases 1 to 348)
AUTHORS
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlroge,J., Ralkehl,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL
95148729
MEDLINE
COMMENT
On May 5, 1995 this sequence version replaced gi:797667.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ldm.cl.msu.edu
The sequence entry for this EST has been reverse complemented and
is being submitted in the sense orientation.
Seq primer: M13-21.
FEATURES
source
1. 348
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone="185L1XP"
/clone_1lb="lambda-PRU2"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT
110 a 74 c 66 g 93 t 5 others
ORIGIN
Query Match 32.2%; Score 315.6; DB 42; Length 348;
Best Local Similarity 96.4%; Pred. No. 1e-69;
Matches 321; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

RESULT 10
AM443205
LOCUS
EST308135 tomato mixed elicitor, BRI Lycopersicon esculentum cDNA
DEFINITION
clone CLEF43M10 5', mRNA sequence.
ACCESSION
AM443205
VERSION
AM443205.1 GI:6985387
KEYWORDS
EST.
SOURCE
tomato.
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 608)
AUTHORS
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from tomato callus (mixed elicitor)
Unpublished (1995)
COMMENT
On Dec 20, 1995 this sequence version replaced gi:1135853.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
source
1. 608
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande Fluor"
/db_xref="taxon:4081"
/clone="CLEF43M10"
/clone_1lb="tomato mixed elicitor, BRI"
/tissue-type="leaf"
/dev stage="4-6 week old plants"
/lab host="XLI-Blue MR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenchone, Etx,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
BASE COUNT
165 a 126 c 134 g 183 t
ORIGIN
Query Match 31.9%; Score 312.6; DB 81; Length 608;
Best Local Similarity 71.5%; Pred. No. 6.2e-69;
Matches 411; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Db	273	CTTAAACATCTACAGTATCTGAGAGCTTTACAAAAATATAATATCAGGAAACATCCTTAAG	332
QY	440	gaacttgygaaatctgaaagaatcctcatcagcttggatctgtacaacaacaatcttaacgg	499
Db	333	GAGCTCGTAACTTGAAGAGCGTTATTAAGTCTGGATCTGTACAAACAACAATATTTGGGG	392
QY	500	atagttcccaactcttttgyggaaatgaagtcctctgtctctttaaagcgttaatgcaaac	559
Db	393	ACAATTCCTACTTACTTGTGGAAACCTTAAAAACCTTTTCTTTTGCGCTTAAATGATAAC	452
QY	560	cgattgaccggtcccaatcccttagagcactacagycnaatcccaagccttaagttgtgac	619
Db	453	ANGTTAACAGAGCAACATGCCAAGAAGAACTTACTAGCATTTCTAGCCTGAAGAAGTTGGAT	512
QY	620	gtctcaagcaatgtattgtgtgtgaaacaatcccaacaagaagacccttgcacatcct	679
Db	513	GTCCTCGATATATGATTTGTGTGGAAACAATTCCTCTCTGGTCATTTTGAGCATATTCCT	572
QY	680	ttacagaactttgagaaacaaccggaattggagg	714
Db	573	CTAAACATTTTGAGACACAAATCCTCGACTTGAAG	607

RESULT	11
LOCUS	AM031110
DEFINITION	720 bp mRNA EST 15-SEP-1999
ACCESSION	AM031110
VERSION	AM031110.1
KEYWORDS	GI:5869866
	EST.

ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteroideae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 720)
Alcala, J., Vrebalov, J., White, R., Matero, A. L., Viston, T.,
Holt, I. E., Liang, F., Upton, J., Craven, M. B., Bowman, C. L., Ahn, S.,
Rondani, C. M., Fraser, C. M., Martin, G. B., Tanksley, S. D. and
Giovannoni, J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
JOURNAL
COMMENT
On Dec 20, 1995 this sequence version replaced gi:1133546.

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dirisch@clemson.edu
5 prime sequence.
Location/Qualifiers
1..720

FEATURES	source	Location/Qualifiers
	1..720	/organism="Lycopersicon esculentum"
		/cultivar="RA96"
		/db_xref="taxon:4081"
		/clone="GEC13C21"
		/clone_id="tomato callus, TAMU"
		/tissue_type="callus"
		/dev_stage="25-40 days old"
		/db_host="X1-blue MRF"
		/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni laboratory; clec - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST library"
BASE COUNT	200 a	
ORIGIN	149 c	143 g 228 t

Query Match	31.88;	Score 311.6;	DB 64;	Length 720;
Best Local Similarity	71.4%;	Pred. No. 1.1e-68;		
Matches 410; Conservative	0;	Mismatches 164;	Indels 0;	Gaps 0

Db	147	GGTTGTGGCAGTGTTCCTGCTGTACCTGTGGCTGTAAAGGGAATTGAAAGGGAT	206
QY	200	gcctctaaagccttcgcgcggagtttgacaatccagacaaatgctctccagagctggat	259
Db	207	GGTTGTAGCCCTCCGCCGAGAGCTTACTGACCCGGTAACGTTACAGAGCTGGAT	266
QY	260	ccaactcttgttaacccttgaacctggtctcatgtatgaacctgtaaccaagacaacggctc	319
Db	267	CCAAATCTGTGTAACCCCTGTGACCTGGTTATGTCACTGTCAACGGAGATATCAAGTT	326
QY	320	actcgtgtgattgttggaatccaacctcctgagacatcttgccctgagcttggaaag	379
Db	327	ACCGGTGTGATCTTGGGAACCTCAAGTATCTGGTATTGTGTACCTGAGCTCGAAG	386
QY	380	cttgaacattacagtatcttagagctctcaaaaaaacaatccaagaactatccctcc	439
Db	387	CTTGAACATCTACAGATCTGGAGCTTTACAAAATAATATTACGGGAACATCCCTAAG	446
QY	440	gaacttggaaatcctgaagatctcaoaagcttgatctgtacaacaacaatcttacagg	499
Db	447	GAGCTGGGTACTGTGAAGAGCTTTTGTGTGATCTGTCAACAACAATATTTCGGGG	506
QY	500	atagttccacactcctttggygaaatgaagctcctgctctttaaaggctaatgacaac	559
Db	507	ACATTTCTCACTTCACTTGGAAACCTGAAAAACCTTGTTTTCTTGCGCTTAAAGATAAC	566
QY	560	cgaattacccggtccaatccctctagagcactcaaggaacatccaagacccctaaagtgtgac	619
Db	567	AACTTAACAGGACCAATCCCAAGAGAACTTACTACATTTTCAAGCTCAAAAGTTGTGAT	626
QY	620	gtctcaagaatgattgtgtggaacaatcccaacaagaacggaccccttgcacattcct	679
Db	627	GTTCAAAATATGATTGTGTGGAAACAATTCCTACTTCTGTGCATTTGACACATATTCCT	686
QY	680	ttacagaactttgagaacaacccgagattgggg	713
Db	687	CTAAACATTTTGAGCACAATCTCTGACTTGAAG	720

RESULT	12
LOCUS	Z34606
DEFINITION	234606 Versailles- VB Arabidopsis thaliana cDNA clone VBV0H02,
ACCESSION	337 bp mRNA EST 11-JUL-1994
VERSION	ARTS3366 Versailles- VB Arabidopsis thaliana cDNA clone VBV0H02,
KEYWORDS	mRNA sequence.
SOURCE	Z34606.1 GI:507009
ORGANISM	EST. thale cress.
REFERENCE	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 337)
AUTHORS	CNRS.
TITLE	The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL	Unpublished (1996)
COMMENT	Contact: Despeyz T., Amsellem J., Chiapello H., Rouze P., Caboche M., Hofte H.
FEATURES	INRA Versailles Laboratoire de Biologie Cellulaire Route de Saint-Cyr, 78026 Versailles Cedex, France Email: thieryversailles@inra.fr.
SOURCE	Location/Qualifiers 1..337

/organism="Arabidopsis thaliana"
 /strain="ecotype Columbia"
 /db_xref="taxon:3702"
 /clone="BYVDH02"
 /clone_lib="Versailles-VB"
 /tissue_type="whole seedlings"
 /dev_stage="in vitro-grown etiolated seedlings, 5 days old"
 /note="Vector: pBluescript"
 /note="Vector: pBluescript"

BASE COUNT 73 a 98 c 60 g 105 t 1 others

ORIGIN

Query Match 31.1%; Score 305; DB 20; Length 337;
 Best Local Similarity 97.6%; Pred. No. Se-67;
 Matches 330; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

17 ttgtcttccctctgttcagaaatttcccttacttccaaattccttgcatt 76
 |||||
 Db 1 TTGCTTTCCTCTTGTTCAGAAATTTCCCTTACTTCGAATTCCTTTGATT 60
 |||||
 QY 77 cccctcttaacctccgaagctcacatggcgtctgaactatcggtggagctcttc 136
 |||||
 Db 61 CCCCTCTTAACTCCGAAAGCTCAATGGCTCTCGAAACTATCGTGGAGCTCTTC 120
 |||||
 QY 137 gcaagcttcgttaacctgaacttagcttgatcaccttgctcgaagcaactccgaagga 196
 |||||
 Db 121 GCAGCTTCGTTAACCTCACTTACCTTTCATTCACCTGTCGAACCAACTCCGAA- GA 179
 |||||
 QY 197 gatgtctctacagctcttcgcggaatttgacagatccagacatgctctccagaagcttg 256
 |||||
 Db 180 ATGCTNCTCAAGCTCTTCGCCGAGATTGACATGCAACACCATGTCTCCAGAGCTGG 239
 |||||
 QY 257 gatccaaatttgttaacctgttacctgttcacatgtaacctgttaacaaagaaacgcgc 316
 |||||
 Db 240 GATCAACTCTTGTATTACCTTGACTGCTGATGCTCAATGCTGTAACCAAGCAACAGCC 299
 |||||
 QY 317 gtacatcggtg- ggaattgggaattcaaacctctctg 353
 |||||
 Db 300 GTCACCTGCTGGGATTGGGAATTCAAACCTCTCTG 337
 |||||

RESULT 13
 LOCUS A1100679 339 bp mRNA EST 21-AUG-1998
 DEFINITION 33800 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179j11xp 3', mRNA
 sequence.
 ACCESSION A1100679
 VERSION A1100679.1 GI:3449389
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;
 Arabidopsids
 1 (bases 1 to 339)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
 Retzel,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)

REFERENCE
 1 (bases 1 to 339)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
 Retzel,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)

TITLE
 JOURNAL MEDLINE
 COMMENT

On May 5, 1995 this sequence version replaced g1:797664.
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313c@clm.c1.msu.edu
 The sequence entry for this EST has been reverse complemented and

is being submitted in the sense orientation.
 Seq primer: M13-21.
 Location/Qualifiers
 1. 339

/organism="Arabidopsis thaliana"
 /strain="var Columbia"
 /db_xref="taxon:3702"
 /clone="179j11xp"
 /clone_lib="Lambda-PRL2"
 /note="Vector: lambda Zip-Lox; Site.1: Sal; Site.2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques). The vector is BLU's lambda Zip-Lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA."

BASE COUNT 104 a 75 c 66 g 89 t 5 others

ORIGIN

Query Match 30.4%; Score 298.2; DB 42; Length 339;
 Best Local Similarity 94.6%; Pred. No. 2.6e-65;
 Matches 317; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

534 tggctctttaaaggcttaagacaaacgattgacccggtccaa-tccctagaagcactaacg 592
 |||||
 Db 2 TGTCTTTTACGGCTTAATGACAAACCGATTGACGGGCGCAATTCCTTAGAGCACTACT 61
 |||||
 QY 593 gcaatcccaagccttaagttgttgacgtctcaagaacatgtttgttggaacatccca 652
 |||||
 Db 62 GCATTCGCAAGCTTAAATTGTGATNTCTCAACCAATGATTTGTGGAACANTCCCA 121
 |||||
 QY 653 acaagggaccccttgctcacatcccttaccagaacatttgagaaacccgagattggag 712
 |||||
 Db 122 ACAAGGAGACCTTTCTCACAATCTTCACAACTTGAACACCCAGGTTGGAG 181
 |||||
 QY 713 ggaacgaattactggtcttgcaagctcagacactaacctgacactgaaacaaactggaa 772
 |||||
 Db 182 GGAACGGAATTACTCGGCTTGCAAGCTAGCACTACTGCACTGAAATAATTTGGCA 241
 |||||
 QY 773 aacctgaataagaaattggggggtgacctgtgtaagaacacttaccacttccaat 832
 |||||
 Db 242 AANCTGAATAAGAAATTTGGGGGTTGACCTGTGAAGAACACTTACCACCTTAGCAAT 301
 |||||
 QY 833 atcaatctcttatgtaataagatatatattag 867
 |||||
 Db 302 ANCACATCTACTATGTAAATGAATATATATGTAG 336
 |||||

RESULT 14
 LOCUS AM221939 646 bp mRNA EST 07-DEC-1999
 DEFINITION EST298750 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
 clone cLEN6M3, mRNA sequence.
 ACCESSION AM221939
 VERSION AM221939.1 GI:6533623
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Asterales; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 646)

REFERENCE
 Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
 Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
 Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
 Giovannoni,J.
 Generation of ESTs from tomato fruit tissue
 Unpublished (1999)

TITLE
 JOURNAL

Oy	332	ttgggaattcaaacccctctgcgacactcttgcgccttgatcttggaaagcttgaactta	391
Db	557	nnntggtaacttaaccttaaccttaacccatcttgcgacatttggtaacnngaacttgggaagcttgagactcra	538
Oy	392	cagatctagagctctacaanaaanaaatccaaggaactaactcttcgaacttggaaat	451
Db	537	cagatcttggagttgtacaaanaaacaattccaaggaactatnnnnnnggaacttggaaac	478
Oy	452	ctgaagaatctcaccagcttggactctgtacaacaacatcttcaaggaatagttcccaat	511
Db	477	ctgaagaactctagttagcttggactgtgtacacaacaacatattcaggaacacattccacct	418
Oy	512	tctttgggaanaattgaagctcctgtgctctttaaagcttaatsgaacacgattgacogt	571
Db	417	tcattgggggaattggaanaattgtaactgtgtttttagactcaatgacacagactraactgcgc	358
Oy	572	ccaatcccttagagcaactcaaggcaatcccaagccttaagttgttgcgtctcaagcaat	631
Db	357	ccaatttcccaaggaaactgtctgcgtgttcaagccctttaagtagagatgtcttccacaaat	298
Oy	632	gattgtgtggaacaatcccaacaacacggacccttgcacatctgccttaacgaacttt	691
Db	297	gatttatgtggtacatattccctacactcttgcgcatttcagacatatttcattgaattacttt	238
Oy	692	gagaacaaccccgagatttggagggaaccggaattactcgtcttggaaactaagaactaac	751
Db	237	gagatattatcccccgttggaaagctccagagattgttggagactagaatattatgacacaaac	178
Oy	752	tgcacctgaacaactg 768	
Db	177	tgcctggaatracctg 161	

Search completed: June 23, 2000, 19:06:35
Job time: 27428 sec


```

RESULT 1
LOCUS A67821 789 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 26 from Patent WO9743427.
ACCESSION A67821
VERSION A67821.1 GI:4756643
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 789)
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source Location/Qualifiers
1..789
/organism="unidentified"
/db_xref="taxon:32644"
<2..664
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42248.1"
/db_xref="GI:4756644"
/translation="RRPRPNRYRMLFASLITITLALIHVEANSRGLAYALRSIT
DPDRLQSDPILVNPCTWFHYTCNDNRYTRVDLNSNLSGLAPELKLEHLYLE
LYKNIGITIPSELGNLKLISLDLNNLTGIVPTSLGLKSLVFLRLNDRNLTEPI
PRA/LTAIPSLKVVDSNDLCGTFPTNGFPAHPIPLQNEENRPLRGLGLASYDIN
CT"
BASE COUNT 235 a 191 c 156 g 207 t
ORIGIN
Query Match 100.0%; Score 789; DB 5; Length 789;
Best Local Similarity 100.0%; Pred. No. 1.9e-215;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tcgaccacggcgctcggaacatcgtggagctcttcgacgttcgttaacccaac 60
DB 1 tcgaccacggcgctcggaacatcgtggagctcttcgacgttcgttaacccaac 60
QY 61 cttaagcttgaattcacctgctgcgaagcaaacctcgaagagaatgctcttaacgtcttcg 120
DB 61 cttaagcttgaattcacctgctgcgaagcaaacctcgaagagaatgctcttaacgtcttcg 120
QY 121 ccggaggtttaagaagctccggagcaatgcttcgaagagctggagatcctgttaacc 180
DB 121 ccggaggtttaagaagctccggagcaatgcttcgaagagctggagatcctgttaacc 180
QY 181 ttgtaccgtgttcacatgacacgtgaacaaagacaacgcgtcaactcgtgtgaattgag 240
DB 181 ttgtaccgtgttcacatgacacgtgaacaaagacaacgcgtcaactcgtgtgaattgag 240
QY 241 gaattcaaacctctctgacatccttgccctgagcttgggaagcttgaacattacaagta 300
DB 241 gaattcaaacctctctgacatccttgccctgagcttgggaagcttgaacattacaagta 300
QY 301 tctagaagctctcaaaaacaacatccaaggaactaaccttcgaagcttgaagaaatctgaa 360
DB 301 tctagaagctctcaaaaacaacatccaaggaactaaccttcgaagcttgaagaaatctgaa 360
QY 361 gaattcatcagcttgaatctgtacaacaacatcttaacaggaagttcccaattcctt 420
DB 361 gaattcatcagcttgaatctgtacaacaacatcttaacaggaagttcccaattcctt 420
QY 421 gggaaaaatgaagctctctgtcttttaacgcttaatacaacaggaatgaaggggccaat 480
DB 421 gggaaaaatgaagctctctgtcttttaacgcttaatacaacaggaatgaaggggccaat 480
QY 481 ccctagagcaactcactgcaatcccaagccttaagcttgatgtcctaaagcaatgatt 540
DB 481 ccctagagcaactcactgcaatcccaagccttaagcttgatgtcctaaagcaatgatt 540

```

```

QY 541 gtytgaaacaatcccaacaacaggaaccttgcctacatccctttacaagaacttgaaga 600
DB 541 gtytgaaacaatcccaacaacaggaaccttgcctacatccctttacaagaacttgaaga 600
QY 601 caaccgaggttggagggagcgaattactgcgtctgcaagcactgaacactaacgcac 660
DB 601 caaccgaggttggagggagcgaattactgcgtctgcaagcactgaacactaacgcac 660
QY 661 ctgaaaaaatctggcaaaacctgaaatgaagaattgggggggtgaccttgaagaacact 720
DB 661 ctgaaaaaatctggcaaaacctgaaatgaagaattgggggggtgaccttgaagaacact 720
QY 721 caacccttaacatacatcatctactatgtaataagatatataatgtagtgcacaaaaa 780
DB 721 caacccttaacatacatcatctactatgtaataagatatataatgtagtgcacaaaaa 780
QY 781 aaaaaaaaa 789
DB 781 aaaaaaaaa 789
RESULT 2
LOCUS A67825 1063 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 30 from Patent WO9743427.
ACCESSION A67825
VERSION A67825.1 GI:4756647
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1063)
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source Location/Qualifiers
1..1063
/organism="unidentified"
/db_xref="taxon:32644"
106..762
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42252.1"
/db_xref="GI:4756648"
/translation="MASRNYRMLFASLITITLALIHVEANSRGLAYALRSITDP
DPRVLSMDPILVNPCTWFHYTCNDNRYTRVDLNSNLSGLAPELKLEHLYLE
LYKNIGITIPSELGNLKLISLDLNNLTGIVPTSLGLKSLVFLRLNDRNLTEPI
PRA/LTAIPSLKVVDSNDLCGTFPTNGFPAHPIPLQNEENRPLRGLGLASYDINCT"
BASE COUNT 313 a 242 c 206 g 302 t
ORIGIN
Query Match 97.5%; Score 769.6; DB 5; Length 1063;
Best Local Similarity 98.9%; Pred. No. 6.9e-210;
Matches 775; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 5 cccacgctccgggaacatcgtggagctcttcgaggttcgttatcctaactta 64
DB 103 cccacgctccgggaacatcgtggagctcttcgaggttcgttatcctaactta 162
QY 65 gctttatccactggtcgaagcaaacctcgaagagaatgctcttaacgcttcgcgg 124
DB 163 gctttatccactggtcgaagcaaacctcgaagagaatgctcttaacgcttcgcgg 222
QY 125 agttaacagatccggaacatgcttcgaagagctggagatccaactctgttaactctgt 184
DB 223 agttaacagatccggaacatgcttcgaagagctggagatccaactctgttaactctgt 282
QY 185 acctggtcactgtaaccctgtaacaaagacaacgcgtcaactcgtgtgattgggaat 244

```


Db 283 ACCTGTTCCATGTCACCTGTTAAACAGACACCGCGTCACCTGCTGTGATTTGGGGANT 342
Qy 245 tcaaacctctctggacactcttgccctgagcttgaggaaacttgaaacattacagatctca 304
Db 343 TCAAACTCTCTGGACATCTTGGCCGTGAGCTTGGGAGCTTGAACATTATACAGTATCTA 402
Qy 305 gagcttcaaaaaaacaacatccagaagactatctctcgaacttggaattcgaagaat 364
Db 403 GAGCTTACAAAACAAACATCCAGGAACTATCTCCGAATTTGGAAATCTGAAGAAAT 462
Qy 365 ctcatcagcttgatctgtaacaacaacatcttaccagggatgctccactctcttgga 424
Db 463 CTGATCAGCTTGGATCTGTACAAACAACATCTTACAGGGATAGTTCCACTCTTTGGGA 522
Qy 425 aaattgaagctctgctcttcttaagcgttaattgacaacggatgtagcggggccaatccct 484
Db 523 AAATTGAAGCTCTGCTGCTTTTACCGCTTAAAGACACGATGAGGGGCAATCCCT 582
Qy 485 agagcactcaactgcaatcccaagccttaaaagtgttgatgctcgaagcaatgattggt 544
Db 583 AGAGCACTCACTGCAATCCCAAGCCTTAAAGTTGTGATGCTCAAGCAATGATTTGTCT 642
Qy 545 ggaacaatcccaacaagaagacctttgtctacatctcctttacagaacttgagaacaac 604
Db 643 GGAACAAATCCCAACAAACGACCTTTTGTCTACATTCCTTTACAGAACTTTGAGAACAC 702
Qy 605 ccgaggttgaggagcgcggaattactcgtgcttgcaagcttagaagcactaactgcaactga 664
Db 703 CCGAGGTTGGAGGACCGGAAATTAAGTGGCTTGCAGAGCTACACACTTACTGCACTTGA 762
Qy 665 aaaaattggcaaaaacctgaanaattggaagattggggggtgagccttgtaagaacacttcac 724
Db 763 AAAAATTGGCAAAACCTGAAATGAAAGAAATGGGGGTGACCTTGTAAACACTTCACAC 822
Qy 725 acctatcaaatatcacatctactatgtaataagttatataatgtagtccaaaaaataa 784
Db 823 ACTTATCAAAATATCACATCTACTATATTAATGATATATATGATGATCCAAAAAATAA 882
Qy 785 aaaa 788
Db 883 TGA 886

RESULT 3
LOCUS A67819 981 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 24 from Patent WO9743427.
ACCESSION A67819
VERSION A67819.1 GI:4756641
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 981)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICRITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source location/Qualifiers
1..981
/organism="unidentified"
/db_xref="taxon:32644"
104..760
/note="unnamed protein product"
/codon_start=1
/protein_id="CA842246.1"
/db_xref="GI:4756642"
/translation="MASKRWELFPAASLTLLALIHVEANSESDALYALRSITDP
DHVQSDPTLVNPTGWFHYTCQDNRVTRVDLGSNSLGHAPLGLKLEHLYTELY
KNNIQGTIPSELGNLNLISLDYNNNLGIPTSLGKLSLVFRLDNRITGPIPR
ALTAIPSLKVVVDVSSNDLCGTTPTNGPFAHIPILQENFNPRLEGGELGLASYDNCT

BASE COUNT 286 a 236 c 180 g 279 t

ORIGIN
Query Match 94.0%; Score 742; DB 5; Length 981;
Best Local Similarity 96.8%; Pred. No. 5.5e-202;
Matches 757; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 5 cccacgcgtccggaacatctggtggagcctctgcagcttgtaatcctaactta 64
Db 101 CACATGCGCTCTGGAACATATCGGTGGGAGCTCTCCAGCTTGCTTAACTTAACCTTA 160
Qy 65 gcttattacacgttgcaagaacactccgaagagatgctcttcacgtcttcgcgcg 124
Db 161 GCTTGAATTCAGTGTGCAAGAACTCCGAGGAGATGCTCTTACGCTTCTGCGCG 220
Qy 125 agttacagatccggaacatgtctccagagcttggaatccaaactctgttaactctgt 184
Db 221 AGTTTGACAGATCCAGACATGCTCTCCAGAGCTGGATCCAACTCTGTTATCTTGT 280
Qy 185 acctgttccatgtacacctgttaacaaagcaacccgcgtcactcgtgtgattgggaat 244
Db 281 ACCTGTTCCATGTACCTGTAAACAGAACACCGCTGCTGCTGATTTGGAAAT 340
Qy 245 tcaaacctcttgacatcttgccctgagcttgaggaaacttgaaacattacagatctca 304
Db 341 TCAAACTCTCTGGACATCTTGGCCTGAGCTTGGGAACTTGAACATTTACATATCTA 400
Qy 305 gagcttcaaaaaaacaacatccagaagactaactctccgaacttggaactcgaagaat 364
Db 401 GAGCTTACAAAACAAACATCCAAAGAACTATACCTTCGAACTTGAAGAACTGAAGAAAT 460
Qy 365 ctcatcagcttgatctgtacaacaacatcttacaaggaatgattccactctcttgga 424
Db 461 CTGATCAGCTTGGATCTGTAAACAACATCTTACAGGAGATAGTTCCACTCTTTGGGA 520
Qy 425 aaattgaagctctgctcttcttaagcgttaattgacaacggatgtagcggggccaatccct 484
Db 521 AAATTGAAGCTCTGCTGCTTTTACGCTTAATGAAACACGATTCACCGGTCATCCCT 580
Qy 485 agagcactcaactgcaatcccaagccttaaaagtgttgatgctcaagaacatgattggt 544
Db 581 AGAGCACTCACGCGCAATCCCAAGCCTTAAAGTTGTGAGCTCTCAAGCAATGATTTGTCT 640
Qy 545 ggaacaatcccaacaagaagacctttgtctacatctcctttacagaacttgagaacaac 604
Db 641 GGAACAAATCCCAACAAACGACCTTTGCTCACATTCCTTTACAGAACTTTGAGAACAC 700
Qy 605 ccgaggttgaggagcgcggaattactcgtcttggaagcttagaagcaactggaactga 664
Db 701 CCGAGATTGGAGGACCGGAAATTAAGTGGCTTGCAGAGCTACGACATACTGCACTGGA 760
Qy 665 aaaaattggcaaaaacctgaanaattggaagattggggggtgaccttgtaagaacacttcac 724
Db 761 AACAACTGGCAAAACCTGAAATGAAAGAAATGGGGGTGACCTTTAGAACACACTTCAC 820
Qy 725 acctatcaaatatcacatctactatgtaataagttatataatgtagcctcaaaaaaataa 784
Db 821 ACTTATCAAAATATCACATCTACTATATTAATGATATATATGATGATGATGATGATGAT 880
Qy 785 aa 786
Db 881 AA 882

RESULT 4
LOCUS A67817 1106 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 22 from Patent WO9743427.
ACCESSION A67817
VERSION A67817.1 GI:4756639
KEYWORDS
SOURCE .
ORGANISM unidentified.

REFERENCE
AUTHORS
TITLE
JOURNAL

unclassified.
1 (bases 1 to 1106)
De V.S., Schmidt E.D., Van, H.G. and Hecht, V.F.
PRODUCTION OF APOICRITIC SEED
Patent: WO 9743427-A 20-NOV-1997;

BASE COUNT	331 a	258 c	206 g	311 t
ORIGIN				

[illegible]

BASE COUNT	270 a	163 c	176 g	285 t
ORIGIN				

QY	361	gaactcatcagcttgatgattgtacaacaaaccttcaagggaagtccaccattcctt	420
Db	375	GAGCTATACGATTGGATGTGTCACAACAATTCACCGGGAAATCCCATCTTCTTT	434
QY	421	gggaaaattgaagtcctcgcttttttacgcgttaattgacaacggattgacggggccaat	480
Db	435	GGGAAAATTGAAGCACTGTGTTTTTGCGCTTAACGAAAAACCGAATGACCAGCTCTAT	494
QY	481	ccttagagacactcctcctaaccgaaccttaagaattgtgatgctcagaacgatattt	540
Db	495	TCTTAGAGACTCAGATTTATTTCAAGCTTAAGCTTGAAGTGTCTCAGGAAATATT	554
QY	541	gtgtggaacaatccccaaaaacggaccctttgtctcacattccttcagaacttggaaa	600
Db	555	GTGTGGAACAATTCACATAGAGAGACCTTTTGAACACATTCATCAAAACTTTGAGA	614
QY	601	caaccggaggtttggaagggaacgggaattacogtctctgcaagcttgaacatacgaac	660
Db	615	CAACCTGAGATTGAGAGGACCAAGACACTAGTGCTTGCGAGCTATGACCACTTGCA	674
QY	661	ctgaaaaaatgycnaaacctgaaatgaaatgggggggagaccttgttaaagaacctt	720
Db	675	TTAAAAGAAAGTTATAAACCTATTAAGAAAGAA-TGTTAGTGACCTTGTAAGACTCTG	733
QY	721	caccaacttatcaatatc	739
Db	734	TACCAAGCTTGTGTAATC	752
RESULT	6	A67827	2089 bp DNA PAT 05-MAY-1999
LOCUS	A67827	Sequence 32 from Patent WO9743427.	
DEFINITION	A67827		
ACCESSION	A67827		
KEYWORDS	A67827.1	GI:4756649	
SOURCE			
ORGANISM		thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 2089)		
AUTHORS	De V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.		
TITLE	PRODUCTION OF AFOMICTIC SEED		
JOURNAL	Patent: WO 9743427-A-20-NOV-1997;		
CIBA GEIGY AG (CH)			
FEATURES	Location/Qualifiers		
source	1..2089	/organism="Arabidopsis thaliana"	
		/db_xref="taxon:3702"	
		/clone="SPRK GENE CDNA"	
	195..2072	/note="unnamed protein product"	
		/codon_start=-1	
		/protein_id="Cab42254.1"	
		/db_xref="GI:4756650"	
		/translation="MESSYVFILLSILLPNHSLWLANLEGDLHTLRVTLPDNP	
		NVLSDMPPLVNPCFPHVCNNENSVIRDDNALSCHIVDELGVKNLYLPS	
		NNITGPSPNLGNLTJNVSDIYLNSRSGIPPSLKIKRLRIANNNSLGTSIPMS	
		LNIITTLQVLDLSNRKSGVPONGSFLSTPSEFNNDLCGPVTSHPESPFPBP	
		PPTLOPVYSISPGIGITAIAGVAAGALFFAAPALAFAMWRKRSLDLEFDPA	
		EEDPEGLAQOLKRFSLAEQLVASDGSNNKNIIDRGFGVYGRADGLIVAKRKE	
		ERTPGSELQOTEVEMISMAVHRNLRLNGFCMTPERLLVYPYANGSVACLBERP	
		PSOPGLPPTPRKIALGASARGLSYLHDHCPIKHIDVKAANILDEFEAVDGEL	
		AKLMDDYDTHTTAIVRGTIIGHIAPEYLSSTSEKSEKDVGYGIMLELTGRRAPLA	
		RLAADDVMLDMWKGLKREKLMEVDLPDIQNYERELEOVIOVALCTOGSPMER	
		PKMSEVVRMLEGDLAKMKEMOKVELIREIDLSRPNSDWILDSTYNLHAVELSGP	
BASE COUNT	569 a	419 c	502 g 599 t
ORIGIN			

Query Match 26.8%; Score 211.8; DB 5; Length 2089;
Best Local Similarity 63.0%; Pred No.4,6e-50;
Matches 327; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY	86	gcaactccgaaggagatgctcttaccgctcttcgcgcggaggttaacagatccggacct	145
Db	270	GCTAATTTGGAAAGGTGATTCCTTGACACTTATGAGGGTTACTGTAGTGCACAAACAT	329
QY	146	gtttccagaagcttbgagatccaaactctgtttaacccttfgaacctggtttcatcactgt	205
Db	330	GTCCTTGCAAGACTGGATCCTTAGCTAGTGAATCCTTGCAATGATGTTCCATGTCTTC	389
QY	206	aaccaagaacaacccgcgcacacgcctgctgctgctgctgctgctgctgctgctgct	265
Db	390	AACAACGAGAACAGTGTCAATAGAGTGTGATTTGGGAATGCAAGATTAATCTGGCATTTA	449
QY	266	gcgccttagcttggagagcttgaacattcagatcagatcagatcagatcagatcagatc	325
Db	450	GTTCCAGAGCTGTGCTGCTGCCAAGATTTGCAATTTTGAGCTTTACAGTAAACAATA	509
QY	326	caaggaaatbacctccgaacttggaaatctgaaatctcaatctcagcttgcctgtac	385
Db	510	ACTGGCCGATTCCTAGTAATCTTGGAATCTGACAAACTTAAGTATGATTTGATCTTTAC	569
QY	386	aacaaactcttaacaggatgattgccactcttcttggagaaatctgaagctcgtgcctt	445
Db	570	TTAAACAGCTTCTCCGGTCTCATTTCCGGAAATCTATGGGAAAGCTTTCAAAGCTGAAAT	629
QY	446	ttaaggcttaatgacaacacgattgacggggccaatccctcagagcaactcagatccaca	505
Db	630	CTCCGGCTTAACAACAACAGTCTCACAGGCTCATCTGATTCATGCTACGACCATATTA	689
QY	506	agccttaagcttggctgtctcagaagaatgattggtgagcaatcccaacaacagca	565
Db	690	AACCTCAAGTGTATGATCTATCAAAATACAGACTCTCTGCTCAATCTCTGACATRGSC	749
QY	566	ctttgtctacatctccttaacagaacttggaaacac	604
Db	750	TCTTCTTACTCTTACACACCACCATGCTTTGCTAATMAC	788
RESULT	7	SBU62279	936 bp mRNA PLN 09-DEC-1996
LOCUS	SBU62279	Sorghum bicolor leucine-rich repeat-containing extracellular	
DEFINITION		glycoprotein mRNA, complete cds.	
ACCESSION	U62279		
VERSION	U62279.1	GI:1710123	
KEYWORDS			
SOURCE		sorghum.	
ORGANISM		Sorghum bicolor Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.	
REFERENCE	1 (bases 1 to 936)		
AUTHORS	Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.		
TITLE	Isolation of a cDNA encoding a novel leucine-rich repeat motif from		
JOURNAL	Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)		
MEDLINE	97124217		
REMARK	Erratum: [[published erratum appears in Mol Plant Microbe Interact		
	1997 Mar;10(2):302]]		
	2 (bases 1 to 936)		
	Hipskind,J.D.		
	Direct Submission		
	Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant		
	Pathology, Purdue University, West Lafayette, IN 47907, USA		
FEATURES	Location/Qualifiers		
source	1..936	/organism="Sorghum bicolor"	
		/db_xref="taxon:4558"	
		/clone="SLRR"	
	12..755		
CDS			

```

/function="protein-protein ligand binding"
/notes="leucine-rich repeat-containing extracellular
glycoprotein; contains six N-glycosylation sites
[NX(S/T)]"
/codon_start=1
/protein_id="AAC49559.1"
/db_xref="GI:1710124"
/translation="MAPQFAAGFLGLATLAFASQNTGDIILYKORLAMEDPNVYL
QSNRNATLNCCTWEHYTCNNNNFYIRIDNAGISGLPDLAIONLOYELXGSL
NGSIPETLGNLNLISLDMDNLITGPTLVSSTRLRIRYKNNLTGPIPSFGN
LTSLSKLQENLSGALPASLGNLKAQSRSLNDNNLTGTPSKSPSTFSGNLT
NTRNNLMDGTRTSSGSLKVTAAIIDALAKTA"
36..77
e19.68; Score 155; DB 8; Length 936;
Best Local Similarity 58.2%; Pred. No. 8.2e-34;
Matches 272; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

s19_peptide /evidence=not experimental
BASE COUNT 247 a 232 c 207 g 250 t
ORIGIN

Query Match 19.68; Score 155; DB 8; Length 936;
Best Local Similarity 58.2%; Pred. No. 8.2e-34;
Matches 272; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

Db 81 AACCTGAAGGTGCATCATCTGTACAAAGCAAAAGTTGGCATGGAGGACCCAAACACGTG 140
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 149 ctccgaagctggagatccaaactctgttaactctgttacctgtgtccatgtaacctgtaac 208
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 CTGCAGACCTGGAAATTCGACGCTGCCAATCCCTGCCACCTGTTCCATGTCACCTGCAGAC 200
Qy 209 caagacaacgcgcgaactcgtgtgtgattgtgggaattcaaacctctctgtgacatctgcg 268
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 201 AACCATTAATTGTGCATCCGGTGGATTGGGCAATCGAGGCATCTCCGGTCTCTGCTT 260
Qy 269 cctgaagcttggaaagcttggaaacatttacagatctagaagccttacaacaacaacaccaa 328
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 CCAGATCTTCGAGAAATTCAGAACCTCCAGTACATCGACGTGTATGGCAACGGCTTGAAAC 320
Qy 329 ggaactatacctccgaacttggaaatctgaaatcagaatcctcaatcagcttggatgtgtacaa 388
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 321 GGTTCGATTCAGAAACACTAGGCAACCTGCATATCATCAACGCTTGAGATCTCGGGAC 380
Qy 389 aacaatcttaagagatagtcctccactctcttggaaaatggaagtctctgctcttta 448
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 AACCTCTTACCGGCAAAATCCCAACATCGCTTGCTTCTGACGACGCTGCGATATCTG 440
Qy 449 cgccttaatgacaacagatgacggyggacaatcccttaagacatactgcaatcccaagc 508
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 441 AGGTGTATCCAAACAAACACTGCAGAGGGCTTATACATCATCGTTTGGCAACCTGACTAGC 500
Qy 509 cttaagctgtgtgagtctcaagcaatgatattgtgtggaacaatccc 555
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 501 CTTCTGGAATCGAAAGCTTCAAGAAATTCGTTGAGAGCGGCGCTATTCTC 547

RESULT 8
LOCUS DCU93048 1755 bp mRNA PLN 28-JUN-1997
DEFINITION Daucus carota somatic embryogenesis receptor-like kinase mRNA,
complete cds.
ACCESSION U93048
VERSION U93048.1 GI:2224910
KEYWORDS
SOURCE .
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asterales; euasterids II; Apiales; Apiales; Daucus.
REFERENCE 1 (bases 1 to 1755)
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
TITLE A leucine-rich repeat containing receptor-like kinase marks somatic
plant cells competent to form embryos
JOURNAL Development 124 (10), 2049-2062 (1997)

```

FEATURES	source
1. 1755	/organism="Daucus carota"
	/db_xref="taxon:4039"
34. .165	/note="SERK"
	/codon_start=1
	/product="somatic embryogenesis receptor-like kinase"
	/protein_id="AAB61708.1"
	/db_xref="GI:2224911"
	/translation="MNRSINILNYMOTFAYLDKCYVMTLEYSNNISGPIPSDLNITNLSIDLITNFSQFIPIPTDLKLRFLRNLNNSLSGPIFMSLTNITITQVLDNNRISGVPDNGSFSLETFPIFSANNLNLCGPATGRRCSPSPSPPIPTSTVOPCONPSTGAIGVAAAGALALFLFAPAAAFMMRRKRREHFPPVPEEPEVHLGQLRSLSLQVATIDTSTILGRGFEKVKYKGLAGSLVAVRLKEEPTPGSELQFTEKMSIAVNRILRLRGFCMPETERLLVYPYMANGVNSVCRLREOPESEPLADPTRRALGSAAGLSYLDHODKPLIHRQVKANIIILDEEFEEVAGDFELALMDYKDHVHTVAGTIGIYAPETLSGKSEKTDYFGCINILYELINQRAFPDLARLNDVWLLDMKSLKLEKLEMDLVDLENNITIDTEVQLLYALCLTQGSPPMRPKSEVRYRLBEDIAEKEDWQKVEVYHODELAPHTFTSSWILDSTDNCLAFELSGPR"
BASE COUNT	506 a 347 c 407 g 495 t
ORIGIN	
Query Match	13.1%; Score 103.4; DB: 8; Length 1755;
Best Local Similarity	55.4%; Pred. No. 5.2e-19;
Matches 200; Conservative	0; Mismatches 161; Indels 0; Gaps 0
QY	265 tgcgcctgagctggggaagctgaacattacagatctagaagctctacaacaaacacat 324
DB	78 TGAGCTTACCTTGACAAATATGGGGTCTTTAGCAATTGGAGCTTTACACAAATPACAT 137
QY	325 ccaaggaactataccttcgcgaacttggaaattgaaagaatctcaatcagcttgatcgta 384
DB	138 AAGTGGACCAATTCCTAGTGAATCTTGGAATCTGCAAAATTGGTGACCTTGGACCTATA 197
QY	385 caaacaacatcttacaggagtagtctccaactcttgggaaatggaagtctcgctct 444
DB	198 CATAAATAGCTTCTCTGGACCTTACCGGACACATTAGGAAGCTTACAAGCTTAAGATT 257
QY	445 ttacagcttaatgacacagcattgacggggccaatccctagagcactacatgcaatcc 504
DB	258 CTTGCGCTCAACAACACAGCCTCTCTGCTCAATTCGAATGTCACGTCAATTAATTATAC 317
QY	505 aagcctaaagttgttgatgctcgaagcagatattgttggtaacaatcccaaaaagg 564
DB	318 AACTCTCAAGTCTCTGGAATTATCAACAATGGCGCTATAGAGCCAGTACCGGATAATGG 377
QY	565 acccttgcacatctcctttacagaacttgaagaacaaccgaaggttggaggagccgga 624
DB	378 CTCATTTTCTTTGTTTACACCTACGTTTTCGCAATATATTGAATTAATGAGCCCGT 437
QY	625 a 625
DB	438 A 438
RESULT	9
LOCUS	A67797 1815 bp DNA PAT 05-MAY-1999
DEFINITION	Sequence 2 from Patent WO97/43427.
ACCESSION	A67797
VERSION	A67797.1 GI:4756623
KEYWORDS	
SOURCE	carrot.
ORGANISM	Daucus carota
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Matches	120;	Conservative	0;	Mismatches	35;	Indels	0;	Gaps	0;
QY	80	gtcgaagcaactcgcgaagagatctcttaacgctcttcgcgcggagttacaagatcgc	139						
Db	194	GTAAAGGGGAAATTCACAAAGGGGATGTTGTATGACGGCTTCGCGAGACTTATATTCGACCCG	233						
QY	140	gaccatcttcctcagacgcgcggatcccaactctctgttaacctcttacctggttccatcgc	199						
Db	254	GGRACACTGTTCACAGAGCTGGGATCCAAATCTTGTTAACTCTTGACTCTGGTTTCATCTC	313						
QY	200	acctgttaaccaagacacacgcgcgtacatcgtgtgga	234						
Db	314	ACTTGCAACGAGATATCATCAAGTTACTCTGTGTGA	348						
RESULT	11								
LOCUS	ATAC009991/c								
DEFINITION	Arabidopsis 101284 bp	DNA	PLN	08-OCT-1999					
ACCESSION	Arabidopsis thaliana chromosome III BAC F9f8								
VERSION	AC009991								
KEYWORDS	complete sequence.								
SOURCE	AC009991.3	GI:6016677							
ORGANISM	HTG.								
REFERENCE	Arabidopsis thaliana								
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis								
TITLE	1 (bases 1 to 101284)								
REFERENCE	Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B., Romling,C.M., Koo,H., Fujii,C.Y., Uteerback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.								
AUTHORS	Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence								
TITLE	Unpublished								
REFERENCE	2 (bases 1 to 101284)								
AUTHORS	Lin,X. and Kaul,S.								
TITLE	Direct Submission								
REFERENCE	Submitted (09-SEP-1999) The Institute for Genomic Research, 9712								
AUTHORS	Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org								
TITLE	3 (bases 1 to 101284)								
REFERENCE	Lin,X.								
AUTHORS	Direct Submission								
TITLE	Submitted (08-OCT-1999) The Institute for Genomic Research, 9712								
REFERENCE	Medical Center Dr., Rockville, MD 20850, USA								
COMMENT	On Oct 8, 1999 this sequence version replaced gi:5902413.								
	Address all correspondence to:								
	Xiaoying Lin								
	The Institute for Genomic Research								
	9712 Medical Center Dr.								
	Rockville, MD 20850, USA								
	e-mail: xlin@tigr.org								
	BAC clone F9f8 is from Arabidopsis chromosome III and is near the								
	molecular marker g4547.								
	The orientation of the sequence is from SP6 to T7 end of the BAC								
	clone.								
	Genes were identified by a combination of three methods: Gene								
	Prediction Programs including GRATU (available by anonymous ftp								
	from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of								
	Washington), GENSCAN (Chris Burge,								
	http://genome.stanford.edu/~chris/GENSCANW.html), and NetPlantGene								
	(http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the								
	complete sequence against a peptide database and the Arabidopsis								
	EST database at TIGR (http://www.tigr.org/cdb/at/est.html).								
	Annotated genes are named to indicate the level of evidence for								
	their annotation. Genes with similarity to other proteins are named								
	after the database hits. Genes without significant peptide								
	similarity but with EST similarity are named as 'unknown' proteins								
	Genes without protein or EST similarity, that are predicted by mor								
	than two gene prediction programs over most of their length are								
	annotated as 'hypothetical' proteins. Genes encoding tRNAs are								
	predicted by tRNAscan-SE (Sean Eddy,								
	http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are								

identified by RepeatMasker (Brian Smith, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRATL are annotated as misc features.

Location/Qualifiers

```
1..101284
/oranism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="11"
/map="q4547"
/clone="F9F8"
77..108
/misc-feature
/note="exon predicted by xgrall, quality excellent"
124..259
/misc-feature
/note="exon predicted by xgrall, quality excellent"
complement(347..408)
/misc-feature
/note="exon predicted by xgrall, quality excellent_shadowexon"
355..594
/misc-feature
/note="exon predicted by xgrall, quality excellent"
672..752
/misc-feature
/note="exon predicted by xgrall, quality excellent"
854..973
/misc-feature
/note="exon predicted by xgrall, quality excellent"
complement(1182..1357)
/misc-feature
/note="exon predicted by xgrall, quality marginal_shadowexon"
complement(1516..1694)
/misc-feature
/note="exon predicted by xgrall, quality excellent_shadowexon"
1546..1666
/misc-feature
/note="exon predicted by xgrall, quality good"
1747..1838
/misc-feature
/note="exon predicted by xgrall, quality excellent"
complement(2090..2250)
/misc-feature
/note="exon predicted by xgrall, quality good_shadowexon"
2114..2246
/misc-feature
/note="exon predicted by xgrall, quality excellent"
complement(2975..3024)
/rpt_family="(CAA)n"
misc-feature
join(3070..3325,3565..3642,4094..4234,4338..4414,4506..4655,4800..5013)
/misc-feature
/gene="P9F8.1"
3070..5013
/misc-feature
/gene="P9F8.1"
/note="similar to unknown protein GB:AAC27167 [Arabidopsis thaliana]"
join(3215..3325,3565..3642,4094..4234,4338..4414,4506..4655,4800..5013)
/misc-feature
/gene="P9F8.1"
/note="unknown protein"
/codon_start=1
/protein_id="AAFO1505.1"
/db_xref="GI:6016678"
/translocation="MWGAPPOIVFGSSIVOMRSGHGCGAIISEVYARRAKDITLRRCYGNMSRAIEVVDQVPRDAVOPSLIVYEGDMSAPSSGIGPVPLEIYVDNNAISGPIYKRIAHADISDPFRIIFRISPPNDPEAVRONOSPLSEVIRTDNICKYDACEVLEDELGLAEVVDLESTFOKADDMKIVCTDGIHLSAOSKTYVAGELIRVKAELMHPSLHKMPTEFADDSPTDVLVSDGQTVNSSSEMTYFWERQWP"
complement(5513..5597)
/rpt_family="(GA)n"
complement(5527..5597)
/misc-feature
/note="exon predicted by xgrall, quality excellent_shadowexon"
join(5589..5795,5897..5929,6013..6235,7264..7387,7467..7932)
/misc-feature
/gene="P9F8.2"
5589..7932
/misc-feature
/gene="P9F8.2"
/note="similar to nucleic acid binding protein Alfin-1 GB:AAA20093 [Medicago sativa]"
join(5696..5795,5897..5929,6013..6235,7264..7387,
```


REFERENCE	gene
4 (bases 1 to 116944) Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altati,H., Araujo,R., Huizar,L., Rowley,D., Brooks,S., Biehler,E., Chao,Q., Dunn,P., Gonzalez,A., Khan,S., Kremetska,I., Kim,C., Lenz,C., Li,J., Liu,S., Luoro,S., Schwartz,J., Shin,P., Toriumi,M., Vytotskaia,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.M. Direct Submission Submitted (18-SEP-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA	gene CDS
COMMENT FEATURES source	gene CDS
On Jul 15, 1999 this sequence version replaced gi:5441915. Location/Qualifiers 1..116944 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" /clone="F3N23" /complement(133..2966) /gene="F3N23.1" /complement(join(133..2271,2355..2966)) /note="81 identical to threonine synthase [Arabidopsis thaliana] (gi14914408). Location of ests GBGF589 3' (gb F15267) and GBG6175 (gb 226034)" /codon_start=1 /product="Putative threonine synthase" /protein_id="AAd55628.1" /db_xref="GI:5903070" /translation="MASFLPSATYFPSSHSELSLPHSAASFVTCVTSASPAVPPQT POKRRSPENIDEARRRHOLNLSARVPNAPPSSTESYSLDEIVYRSQGLL DNOHDFALKRYDEEFMRNLFDSRQKTPMPSGVMSKREWLPEIDDDIVASFG NSLFAERGRKOYLDNDLWYHCGISHRSGKDIQSGVLSQVNRLEKMRPIYIG GCASTGTSALSHYCSAGISTPSTVFLPADKISMAQVPIANGAFVLSITDIPGCM HLIREVTELPITLANSLSRLRGQTALETLLQFNQVDPWVLPVPGNGNTAF YKGFHMCKELVDRIPIRLVCAQANANPLHYKSGFKEDFNPILKANTFSAIIG DPSIDRAVYALKRNSIGVEATEEELMDATLADSTGMEICPTGTGALTALMKLKS GVGANDRTYVVSAAHGLKFTQSKIDYHSKNIKEMACRLANPVVKAKFGSMDVLK EYLKSNDR" 6030..7079 /gene="F3N23.2" 6030..7079 /gene="F3N23.2" /note="Unknown protein; location of ests 205D12T7 (gb H77203) and 205D12XP 3' (gb AA605559)" /codon_start=1 /protein_id="AAd55629.1" /db_xref="GI:5903071" /translation="MNLGAAEESAOEIHLPADINWMLDKSFVYLGALFSGSAGA LYPAVLKTRQVCHSQSGCIKTAFTLVREHGLRGYRGFTSLMGITIPARALYMAL EYKSNVNGSAVSLGTEAKAAVANAAGLSAAMAQLVWTPVAVSRLMVGSGAG LVNARCNVYNGFAPFRKIYRADPGKGLYSGFSLITVAPSNAMWASVAVRW GGICGVYCKRDESGNSTMTKDSKIMAVGVSAIAGVSALITMDPLDKITKTLQ VLDEDSNNKRSRPSIGQTVRNLYREGGWTACTRGIGPCASMSASATMTTITTEFL KRLSAKHHDGYSKS" complement(8272..9661) /gene="F3N23.3" complement(join(8272..8658,8746..8916,9005..9079, 9161..9252,9364..9661)) /gene="F3N23.3" /note="998 identical to Transcription factor [Arabidopsis thaliana] (gi12398525)". /codon_start=1 /product="Transcription Factor" /protein_id="AAd55630.1" /db_xref="GI:5903072" /translation="MMHOMLKKDSATHSPLTPINTISMGVYPTDSVANRBSASL SLKVDSPRGHIGTKQKISFDQDSSSTQSGSTTEVASSGDNDPSKISFSKSSSE ITQRGFSNPKQSGMTGFPIHFAPAQANFSFADPHYGGLAATYVQAQPTCPQ MVSMPGVPLPAELTETDPVYNAKQYHAIMRRQQRALKEAQNKILIRAKKYLIES RHVAHLKRRSGGRFLNWKTLQESQAAAREODKLGCOVNRKTKNSRFFAHMIO NNKRSSTSSSDITSVDGADIDGHEFQPSGPTPINNAALVHQSNDMHGGDMH HPSVHI"	gene CDS
complement(12076..16057) /gene="F3N23.4" complement(join(12076..12519,12613..13851,13955..14140, 14368..15451,15672..16057)) /gene="F3N23.4" /note="Similar to disease resistance proteins RPP1-WSB (gi13860165), RPP1-Wsc (gi13860167) and RPP1-WSA (gi13860163) [Arabidopsis thaliana]." /codon_start=1 /product="Similar to disease resistance proteins" /protein_id="AAd55631.1" /db_xref="GI:5903073" /translation="MASSSSSSTRRLRHVDVFLSPRGVDTROTIVSHYVALRNGVY TPRDKRELEGGRIADGIYKATQTSWFAVILISENVAATSTGLIELRLIMOLHSEOI KVLPIYGVKPSVYRQESGFATPSYDEADMIANVGGISRLRPMKSTDLINMG MEAHMMKMTLLNIGCEDEHYHMGIMGKGIQSTAKLIRFSHQFARFLEVNS KGYDKHQLKELLSHLYDEVDVLMENEGSDIEKRLHQVFPVLDVNDVEQLHG LAKDPWFQSGSRILITTRDKGLNCSGVNNIVYEVCKDLKDALQVFKLAFGRPPS DGEPOLEFRASRLAHGLPSALVAFASHLSAIVADMEDELALLEFPQKNQELIRA SYGDIQDYDKTVFLHVAACFPNGGHLRYRAFKNCDARLNHAKCLVNVISIDGICSM HLLVQGRATVROESDMRSKORFLMDPTEPIHYVDSMTNHLGANSNQLISDDYVL SRNKLHMDVPLITLPIFRPHTITELSRKSNLSMDGTCLLPNLRITDYGSR NLRELPLSTAVNLEULLESCTSLVQIDPESINRLRLNMYCDGLGVLVDVQ EASLSRWGLKRIILNLPHGATLSYDLDAIQKIRIKLDSGLGTDHLSFSVQTA HOSVTHLNGSFGFKSLDIKRPSTRLDPNFCISFADPCLTELKILNLEIDIE DICOLOLLELDGNDVFLPTSMGOLAMKLSLNSCRRLKALPOLSOVERLYSG CYRIGSLMGITLGNRGVNLDPVCYKSGISGLITISVKSAPGRLELLEISLNSKS LVLSSELSEFTKLTITDLSLSEFRRIPTISIRLSMRTILYNCKRITSIDPESL KYLVAHCESELEHNESNSHFNHDSHCISLSECLSDVLPDMNEPISQEAFLVLC ITKYSIASTNNMRISWREPMRILPKIKAPKLVGFVQIMVCCPEFLQDPASYN WDQESRLYILINKPNLYQSEEMEDNNRPFYKMHVLVYQIPTGTIISAEIDVQFES HVOVEGEDEILICGVHGVFLK" 16543..17907 /gene="F3N23.5" join(16543..16961,17058..17907) /gene="F3N23.5" /note="Similar to N-terminal half of downy mildew resistance protein RPP5 [Arabidopsis thaliana] (gi12109275)". /codon_start=1 /product="Similar to part of downy mildew resistance protein RPP5" /protein_id="AAd55632.1" /db_xref="GI:5903074" /translation="MSSSSEFEVFLSECSDEPKTYVSLDRLLEQDKITTFNKDDSF LAESKLAVVYSESYPIVYCLNOLKINYSSEGLSTLPFYFQVDPVNVKQGY LAPROELGEGYDPDKIOEMRVSLTKTLNLPALDSRYMSMEADITLAINELISNR KPLAKGDGLVMDRQKOTYKLLDFAAEVYLLIGTSGGCGTGLTARAYAEIIS NFKVHVPDKAEKICHDODRDLKLLKKGTTGGIDGIDKIKSTFGHKGILYIDCV NIOKLKEIVYLAHFIQGSRYIFVQDRNLIVSGVHALEVOSLRYDEALQFSHA FDQOKPPTSESLSLRAVHISGFLPLTLKILIGSSLKNGDERWEKELDLEGDOEKAI MEITSKRTYRAGKEEDKEKITSFILLSD" complement(19050..23350) /gene="F3N23.6" complement(join(19050..19667,19764..21344,21686..22799, 22887..23350)) /gene="F3N23.6" /note="Similar to entire sequence of downy mildew resistance protein RPP5 [Arabidopsis thaliana] (gi12109275)". /codon_start=1 /product="Similar to downy mildew resistance protein RPP5" /protein_id="AAd55633.1" /db_xref="GI:5903075" /translation="MASSSSSPIMRWYVFLSFRGEDTRKNIVSHLKLQVLDKGVVTFK DOKRLEGSYVHRHOGTSPTDPKQDSSMPKWTGWRALQVIAISLIGKDETCDEAS MIEYVKDISKLLIMQVDSIDVGNAMHREKISPLMSDSENVPMIGKMGIG KTLQPEIKLDFDSGFPAPKCLEVNSIYRGGVSLAEKFLSTTLGSLKKNKGSV KLGPDELKRGCKRVFVLDVNDVDMKNAPQESSWFQPSRIIITTRDKGLNLY GVRYVEYKCDNDNALQLEFQALFKALPSELYEKLISRWMLQGLVPALEAGL FFRRMTSLKEMDDNLGRIFEADEPSEVIELKISVDEGTEETDQVFLHVALFNGEPLR BATITLDGTVAGGLGKLIAKELIETASGYIKMNLVNDQARALVNGESNRHRC	gene CDS

Query Match 10.3%; Score 81.6; DB 8; Length 116944;
Best Local Similarity 49.5%; Pred. No. 1,3e-12;
Matches 210; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

RGVLMNPEYIELLRKNTPODVEIILHRSNLSFWKEIVKALNSRLITMYLVLN
ILALFFVEVYAGMPNLRDLSDSNLEQLDLSAVLMEELITOGKRLKIPES
ISYLRLLTLDVSYCEELASYITIREINSGROALAYFSKEVETRSINLSGNIH
IOMFLDGNVHSTETEOGPDRLTKREKQAPGELTKREOQOEPKRTIILGSGH
MRGKRYATSEFLDHMMOROLAPDNOALFSTRTOALQPLPHGOSYVKA
QKRSPTKRFHGTSTVDSINFRISSDASFLCSLMFPCKELIILNLIKIPDV
CGKFLKLDMSGNDFFETLPETNQLPRLYASRNCRLKALPALVOLTETKLSGI
NLQSLLEIYAEODGCRQWELWDCSKIRSLDRLHRIKLSYLDLSHFEKLP
SISYELSLRTCLNKCKKLSIEGLPLKLSYASGCEILEVSLPHNSHEDLS
HCFGLRDEHLIAQFLNENEEELGFAFPGPEVSYDFHIDKGSLLIDLPQEW
SPKLGFDACVIVACERPFDIQSPSPYDMWGERFCYILKDFHSDPSDEDEE

QY 165 caactctgttaactctgttacctgttccatgtcaactgttaaccaagccgctca 224
DB 84523 CATCTGAAGCAGCTCATCTTAAGTGTGATATCTGATGATCTTAAGAAGTTA 84582
QY 225 ctctgtgtgatttgggaattcaaacctctctgtgacatcttgccctgttggaagc 284
DB 84583 CGTCTCTCACTTACCGGCTTCAAGGTGTTCAAGCCAAATGGGCTCTGAATAAGGCAC 84642
QY 285 ttgaacattacagatctcagagctctacaacaacaacatccaaaggaactaccccg 344
DB 84643 TCAAAAGCTTAGAGATATTGATATGATGATTTCTCTGGATTTACTTCCA 84702
QY 345 aactgtgaaatctgaagaatctcaatcagcttgatctgtacaacaacatcttaaggga 404
DB 84703 GTTAAAGGAGACTGATGATCTGATGATGATGATGATGATGATGATGATGATG 84762
QY 405 tagtcccaactctctgtgggaataattgaatctctgtcttcttcttcttcttcttcttct 464
DB 84763 AGGTCCCATCTACTCTCGGTACTGATGATGATGATGATGATGATGATGATGATG 84822
QY 465 gattgacggggcccaatccctagagcactcactgtcaatcccaagccttaagtgttgatg 524
DB 84823 CTCTTACTGGGGAATGCTTAATCTTCTTCTGATCCGGTCTGTAATTAATGATG 84882
QY 525 tcccaagcaatgattgtgtgtgaacaatcccaaggaagcacttcttctcaattcctt 584
DB 84883 TTGAGCATTAACACTTAACCGGCTTAATCCCTCAAAATGTTGTGTAAGCATTAAG 84942
QY 585 taca 588
DB 84943 TTCA 84946

RESULT 13

INUT7888 5033 bp DNA PLN 23-NOV-1996
LOCUS IPomoea n11 receptor-like protein kinase (lnrpkl) gene, complete
DEFINITION cds.
VERSION 077888
KEYWORDS U77888.1 GI:1684912
SOURCE Japanese morning glory.
ORGANISM Ipomoea n11
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Convolvulaceae;
Ipomoea.

REFERENCE 1 (bases 1 to 5033)
AUTHORS Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajeevan,M.S.
TITLE Identification and Preliminary Characterization of an Unusual
Leucine-rich Repeat Receptor-like Protein Kinase from Morning Glory
(Ipomoea n11)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5033)
AUTHORS Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajeevan,M.S.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-1996) USDA, ARS, Appalachian Fruit Research

FEATURES
Source Station, 45 Wiltshire Road, Kearneysville, WV 25430, USA
Location/Qualifiers
1. 5033
/organism="Ipomoea n11"
/cultivar="Violet"
/db_xref="taxon:35883"
/tissue_type="young leaf"
237. 251
TATA_signal
TATA_signal
gene
1144. 4566
/gene="lnrpkl"
join(1144, 4108, 4202, 4566)
/gene="lnrpkl"
/note="leucine-rich repeat; protein kinase homolog"
/product="receptor-like protein kinase"
/codon_start=1
/protein_id="AB36558.1"
/db_xref="GI:1684913"

/translation="MRVAVNFTLSICSTSYIAFAALNSDGAALLSRHMTPIPSD
ITQSRNNSDSTPCSMGVEDRQPDITNLISYGISGFGPEISHLKHKKVUSGN
GFESIPISQNGCSLHEDLSNSSTGNIPTDGLQMLNLSLFSNLTGPPESL
LSIPHEVTYFTGNGSLNGSIPSNIGMSELTLMLDNQSGPVSIGNTDGLY
LNDNNLVLPLVTLNMLENLVLDVNNLSLVGAIPDFVSCQIDITISLNNQFTGL
PPGNGCTSLREFAFSCALSGIPSPCFQOLKLDITLYLAGNHFSGRIPLPGCKSM
IDLOLQNOLEGEIPGELGMLSOLOYLHYTNLSEVPLSIWKIOSLOSLQYNNL
SGELPDMTEKLQDLVSLAYENHFTGTVIRGNSLEVLDTLPMFTGHPIPLCS
OKRIKRLLLYNNLEGSVPDLSGCTLERLILENNLNGCLPDEVKONLPLPDS
NFTGPIPLSGNLKNTATITLSNQSISIPEDSLVLEHLNLSHLLGLPSE
LSNCHKSELDASHNLNLSIPSTLSLELRLSGENSGGIPSTLFSNKLNTL
QLGONLADGIPVGAQLQALRSLNLSNKLNLQPLDGLKMLELDELVSNNLSGLT
RYLSTIOSLTFINSHLSEGPVPSLPTFLNSPFSFGNDLQINPADGLAPES
SILPCNMOGNTGKGLSTGLIMYVIGALFLICLFLSARFLPCKSVDELISA
OEGGSLNLNVLEENTLNDXYVIGAGGTIKATLSPDKYAVAKIYFTGIRKGS
SMVEETITGVKVRHRLIKLEFWRILKREYGLITLYMENGSLDILHETNPCKDS
TRHNIATGATGALYHFDGDPRAIYRDIKPNHLLDPLDEPHISDPGAKLLDSAT
SIPSNVQGTGTGAPENAFYIVKRESQVTSIGVALLILTRKRALPSENGEYD
GVWVSVMTOGEGEIQIVDPSLDLDELIDSSVMQVYELSLALRCAKEKEDKRPMDY
VKOLRMSIRSYSSVANKSK"

BASE COUNT 1365 a 1060 c 1017 g 1591 t
ORIGIN

Query Match 10.3%; Score 81.4; DB 8; Length 5033;
Best Local Similarity 49.9%; Pred. No. 1,1e-12;
Matches 205; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 151 ccagagctggatcccaactctgttaacctctgttaccgttccatctcaacgtgaacca 210
DB 1281 CCAGAGCTGGAAATCTTAGATTCACCTCTGTTCAATGCTGGAGCTTAAGATGACAG 1340
QY 211 agacaacccgctactcgtgtgatttgggggaattcaaacctctctgtgacatcttgccc 270
DB 1341 GAGACAAATTTGTGATCTGAACCTCTCCTTAAGAAATCAGGCAATTCGGGGC 1400
QY 271 ttagcttgggaagcttgaacattacagctatcagctctcaaaaacaacatcccaagc 330
DB 1401 CGAAATCTGCATTGAACATTGAAGAGGTGTCTCAGTGGCAATGCTTCTTGG 1460
QY 331 aactatacccttcgaacttgaatctgaagaatcacaacatccagcttggatgtgaacaaa 390
DB 1461 CTCGAATTCCTCCAGTAAGCAATGACGCTCTTGAACATATAGATCTGTCCTCCAA 1520
QY 391 caacttcaaggaagatgtcccaactcttggggaattgaagatcagctctgcttttaag 450
DB 1521 CACCTTACTGTGATATTCCTCGACACCTTGAGCTTGTGAGATTTAAGAACTTAAG 1580
QY 451 gcttaatgacaacccgattgacgggccaatccctagagcaactcactgcaatcccaagcct 510
DB 1581 CCGTCTTTAATTCMCGATGATGCGCCATTTCCTGAGTCTTACTTCAATTCACACTT 1640
QY 511 taaagtgtgtgtctcaaggaatgatttgtgtgaaacaatcccaacaa 561
DB 1641 AGAACTGTTTATTCACGCAATGCTTAATGATGTTCAATCCCTTCAA 1691

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
14	AC021198	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete	AC021198	81513 bp	DNA	PLN	09-FEB-2000
REFERENCE	1	(bases 1 to 81513)					
AUTHORS		Liu, S.X., Sakano, H., Yu, G., Lee, J.M., Lenz, C., Pham, P., Tortum, M., Chio, C., Chio, J., Choi, E., Chung, M., Gonzalez, A., Hwang, B., Liu, A., Vaynsberg, M., Altai, H., Brooks, S., Buehler, E., Chao, O., Conn, L., Conway, A.B., Hansen, N.F., Johnson-Hopson, C., Khan, S., Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R., Federspiel, N.A. and Theologis, A.					
TITLE	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence						
JOURNAL	Unpublished						
REFERENCE	2	(bases 1 to 81513)					
AUTHORS		Theologis, A.					
TITLE	Direct Submission						
JOURNAL	Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA						
REFERENCE	3	(bases 1 to 81513)					
AUTHORS		Theologis, A.					
TITLE	Direct Submission						
JOURNAL	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA						
COMMENT	On Feb 11, 2000 this sequence version replaced g1:6693723. The sequence is of BAC F14D7 from Arabidopsis thaliana chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 6954 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide small overlaps (200 bp) between overlapping submitted clones. The 5' end of this sequence overlaps by 200 bp to the 3' end of the sequence of the clone F1504.						
FEATURES	Location/Qualifiers						
SOURCE	1. 81513						
	/organism="Arabidopsis thaliana"						
	/cultivar="Columbia"						
	/db_xref="taxon:3702"						
	/chromosome="1"						
	/clone="F14D7"						
BASE COUNT	27779 a 13814 c 13868 g 26052 t						
ORIGIN							
Query Match	10.1%; Score 79.8; DB 50; Length 81513;						
Best Local Similarity	53.7%; Pred. No. 4,1e-12;						
Matches 165; Conservative	0; Mismatches 142; Indels 0; Gaps 0						
249	accctctggacatcttggcgccctgagcttgggaagcttgaacatttacaagtatcagc 308						
DB	ATTCTACTGGGGGCATTCGCCGAACTAGGCAACTATGATCCATGTCATTTTAGGT 2555						
QY	tctcaaaaacaacatcatcaaggaacatcatcccttcgaaacttggaaacttgaagaatctca 368						
DB	TGAGTAACAACAACCTTACCGGGTTCATTCCTTCCTTAGGAATCTCAAGAATTTGA 2585						
QY	tcaagcttgagatctgtacaacaacaatcttacaaggaatagttcccaactctcttgggaaat 428						
DB	CTATTCTTATCTTTACGAGAACTATTTAACCTGGTGTCAATTCCTCCCGGAACTAGCAACA 2645						

QY	429	tgaaagctcttgctcttcttaagcgttaataagcaacgcatggaagggcccaatcccaag	488
Db	2646	TGGATTCGATGATGATTGATTACAGTTGATTAACAAACAACACTTACCAGTTCAATTCCTTCT	2705
QY	489	cactcacatgcaatcccaagccttaagtgtgtgactgtcctaagcaatgattgtgtgaa	548
Db	2706	CCATCGGAATCCCAAGAACTTGACGTATCTTTATCTTTACCTGAATATTAATACGTGG	2765
QY	549	caatccc	555
Db	2766	TCATCTCC	2772

RESULT	15	DNA	PLN	08-DEC-1999
LOCUS	ATAC011765	Arabidopsis thaliana chromosome I BAC F1M20 genomic sequence,		
DEFINITION		complete sequence.		
ACCESSION	AC011765			
VERSION	AC011765.4	GI:6539234		
KEYWORDS	HTG.			
SOURCE				
ORGANISM		Arabidopsis thaliana		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE		1 (bases 1 to 134402)		
AUTHORS		Lin,X., Kaul,S., Town,C.D., Beato,M., Creasy,T.H., Haas,B., Rensing,C.M., Koo,H., Fujii,C.Y., Uterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.		
TITLE		Arabidopsis thaliana chromosome I BAC F1M20 genomic sequence		
JOURNAL		Unpublished		
REFERENCE		2 (bases 1 to 134402)		
AUTHORS		Lin,X. and Kaul,S.		
TITLE		Direct Submission		
JOURNAL		Submitted (14-OCT-1999) The Institute for Genomic Research, 9712		
REFERENCE		3 (bases 1 to 134402)		
AUTHORS		Medical Center Dr. Rockville, MD 20850, USA, xlinet@igr.org		
TITLE		lin,x.		
REFERENCE		Direct Submission		
AUTHORS		Submitted (08-DEC-1999) The Institute for Genomic Research, 9712		
TITLE		Medical Center Dr., Rockville, MD 20850, USA		
JOURNAL		On Dec 8, 1999 this sequence version replaced g1:6102640.		
REFERENCE		Address all correspondence to:		
AUTHORS		Xiaoying Lin		
TITLE		The Institute for Genomic Research		
JOURNAL		9712 Medical Center Dr		
REFERENCE		Rockville, MD 20850, USA		
AUTHORS		e-mail: xlinet@igr.org		
TITLE		BAC clone F1M20 is from Arabidopsis chromosome I and is near the		
JOURNAL		molecular marker m1425.		
REFERENCE		The orientation of the sequence is from SP6 to T7 end of the BAC		
AUTHORS		clone.		
TITLE		Genes were identified by a combination of three methods: Gene		
JOURNAL		prediction programs including GRAL (available by anonymous ftp		
REFERENCE		from artur@epm.ornl.gov), GeneFinder (Phil Green, University of		
AUTHORS		Washington), GenScan (Chris Burge,		
TITLE		http://jgpmc.stanford.edu/~chris/GENSCANW.html), and NetPlantGene		
JOURNAL		(http://www.cbs.dtu.dk/net/gene/bsnet/gene.html), searches of the		
REFERENCE		complete sequence against a peptide database and the Arabidopsis		
AUTHORS		EST database at TIGR (http://www.tigr.org/tdb/at/at.html).		
TITLE		Annotated genes are named to indicate the level of evidence for		
JOURNAL		their annotation. Genes with similarity to other proteins are name		
REFERENCE		after the database hits. Genes without significant peptide		
AUTHORS		similarity but with EST similarity are named as 'unknown' proteins		
TITLE		Genes without protein or EST similarity, that are predicted by mor		
JOURNAL		than two gene prediction programs over most of their length are		
REFERENCE		annotated as 'hypothetical' proteins. Genes encoding tRNAs are		
AUTHORS		predicted by tRNAscan-se (Sean Eddy		
TITLE		http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are		
JOURNAL		identified by repeatmasker (Arian Smil,		

genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

FEATURES
source

```
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db_xref="taxon:3702"  
/chromosome="I"  
/map="m1425"  
/clone="FIM20"  
138..269  
/note="exon predicted by xgrail, quality  
marginal_shadowexon"  
complement(229..281)  
/note="exon predicted by xgrail, quality excellent"  
complement(393..453)  
/note="exon predicted by xgrail, quality excellent"  
complement(744..1286)  
/note="exon predicted by xgrail, quality excellent"  
1271..1354  
/note="exon predicted by xgrail, quality good_shadowexon"  
complement(1356..1475)  
/note="exon predicted by xgrail, quality excellent"  
1378..1581  
/note="exon predicted by xgrail, quality good_shadowexon"  
complement(1509..1530)  
/note="exon predicted by xgrail, quality marginal"  
complement(1510<1655..1777,1908..2135,2258..2581,  
2679..2963,3052..3429)  
/gene="FIM20.1"  
complement(<1655..>3429)  
/note="similar to p58 protein kinase GB:AAB59449 [Homo  
sapiens]: contains pfam profile: PF00069 Eukaryotic  
protein kinase domain"  
complement(join(1655..1777,1908..2135,2258..2581,  
2679..2963,3052..3429))  
/gene="FIM20.1"  
/codon_start=1  
/product="putative protein kinase"  
/protein_id="AAFI5905.1"  
/db_xref="GI:6539235"  
/translation="MGCVSKQVSVTPAIDHSQVKNENECSSGGRVYEDPRPI  
LKTIVSRSSKRSQKSGSELSGSGASLSLSTRGNVSRILEACVAGMPAML  
SVNAGEILHGVNLRSDAEKLEKIGQVTSVNFVAVETEGRIVALKVRDFEPE  
SVFMAEELLILRLNHPNLIKLEGLITKSCNQLVETEMHDLTGILSPDIKFT  
TPQIKCMQKLSGLDHCHSRGVMHDIKSGMLLSNEGILLVADGILANSNSGHR  
KKPLKSRVTVLHVRPELLGATDYGASVDMVSGVCEVALLGKPIRGRTEVQLH  
KIFRLGSPEDYWKSKLPHAMLFKQQTYSCLRETLKDLSEINLITLISIDP  
HKRGTAASALVSGYFKTKTPADPSLPIYPSKEIDRHRDEARVYSFTT"  
complement(4069..4131)  
/note="exon predicted by xgrail, quality excellent"  
complement(join(<4608..4760,4885..>4974))  
/gene="FIM20.2"  
complement(<4608..>4974)  
/note="similar to dolichyl-phosphate mannosyltransferase  
dolichol phosphate-mannose synthetase: EMBO J 1998 Sep  
1;17(17):4920-9"  
complement(join(4608..4760,4885..4974))  
/gene="FIM20.2"  
/codon_start=1  
/product="putative dolichyl-phosphate mannosyltransferase  
polypeptide 2"  
/protein_id="AAFI5906.1"  
/db_xref="GI:6539236"  
/translation="MELADAVGLLSISLITFYTFWIIIPVDSDFHRYFL  
PDYALIVPVFAGIALISLISYVIGVMKSKRKA"  
complement(5111..5190)  
/note="exon predicted by xgrail, quality excellent"  
complement(<5342..>7603)  
/gene="FIM20.3"
```

gene

complement(<5342..>7603)

/gene="FIM20.3"

CDS

/note="contains pfam profile: PF01348 Type II intron
maturase"
complement(5342..7603)

CDS

/gene="FIM20.3"

CDS

/product="putative type II intron maturase"
/protein_id="AAFI5907.1"
/db_xref="GI:6539237"

CDS

/translation="MASEKTMFSLAGELASIVSESSSHYDDSKRSPRMEKRSIEL
KIKRYKQCIKNGKSDLLKVIARPEPLRATDCILKNSVNSTENKSGVADSTAE
ELSSGVFVASTTESIVARDKTEVLVLPVSLKVOEALRILEVAFSFKISHS
CHSGGVSALVKTINNISRDMOCFTLSLNKKDLVSFENLISMEKXDSLSILL
RSMFEARYNLNFEFGGPPKGGHLPQEGVLSRYLNIYDRDFHEFYRISMREALGDS
KIDESPCKSLNEMFRQAGOGGLSTGDEVDLRYCCRMDEIYFSGPKVAD
ISESLIGFRNLSLHLDIDETDPSPCESTSLRYGLVLRKNRESPTVAVAKLEK
VRLPLOKEAMTIGTVRIGKKWIGHGKVKSEIETGLADNSTLSQISCHKAGME
TDHWKILIRIMEDVLTSPDRSEERYLSKHVEPVPQELRDATRFQNRKAATVS
STPAULBALPQOSHDPVFFGDVAPATNIGRLRYRGLITRAKARSNSMLILD
TQAILDWTSGLVRWVWYEGCSNFEDEKALIDQIMSCIRTLAAKRIHMEIERK
LDLELSTPSAEDIDQELQHEKLDSPATRDREHLTVGLSNSGLCLSLARLVESEPC
NCFVIGCSMAAPAVYTLHAMERORFPGWKTFGVCIPSSINGRRIGICRQHLDTYGC
QISQAVDFGAWR"

CDS

complement(7885..8010)
/note="exon predicted by xgrail, quality marginal"
8771..8821
/note="exon predicted by xgrail, quality
marginal_shadowexon"
8935..9008
/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

CDS

/note="exon predicted by xgrail, quality
marginal_shadowexon"

OY	231	tgagttctgggaaattcaaacccctcttgacactcttgccctcgagcttgggaacttgaac	290
Db	12043	TAGATTGGGCTACAACAATTTCTGGGACAGTACCACGAAATTTCTCAGATACGAA	12102
OY	291	attacagatctagaagctctcaaaaaaacatccaagaagactataccttcggaacttg	350
Db	12103	GTTTGAAGTCTCTGATTTCTGCTTATATACCTCAGTGGGAAATTCGCCACAGAGATTAG	12162
OY	351	gaactcgaagaatctcatcaactctggagatctgcacaacaacatcttaccaggaatgac	410
Db	12153	GGAAACATGCGGGGGCTTCAAGCAAGCTTGATCTCTCTTTAAACAGCTGACCGGTTGATAC	12222
OY	411	ccaactctcttgggaaaaattggaagctctcggfctctttaagcgcttaatgacaaccgattga	470
Db	12223	CAGGTTCAATTTGGGAAATGTGACCTCTCTTTGTGGCTATACCTTGCAAAACAACTCTCTAT	12282
OY	471	cggggccaactccctagaagcactcactcgaatcccaagccttaaaagtgttgatgctcaca	530
Db	12283	CAGGAGAAATCCCTCGAGGATGTGGCAACTGCACAAAGCCTTTGTGGTTTAAAGTGGCAA	12342
OY	531	gcgaatgattgtgtggaacaatcc	554
Db	12343	ACAACCAAGCTCTCTGGTAGAGATTCC	12366

Search completed: June 24, 2000, 00:21:39
Job time: 46254 sec

[illegible]

Query Match	97.5%	Score 769.6	DB 1	Length 1063
Best Local Similarity	98.9%	Pred. No. 9.8e-23		
Matches 775	Conservative	0	Mismatches 9	Indels 0
QY	5	ccccgcgcgtcggaaacatcgtgtggagcttcgcagcttcgttaatccaaactta	64	
Db	103	CACATGCGCTTCGAAACTATCGTGGAGACCTTCGAGCTCGTTATCTTAACCTTA	162	
QY	65	gcttcgatccacctggtcgaagcaaacctccgaagagatgctcttaacgtcttcgcgg	124	
Db	163	GCTTGATTAACCTCGTGCAGAACCAACTCCGAAGAGATGCTTTACGCTTCGCGCG	222	
QY	125	agtttaacagatccgcagccatgtcttcacagactgtygaatccaactctgttaatccctgt	184	

Dd	223	AGTTTAACAGATCCGGACCAATGTTTCCCAAGCTGGGATCCAACTTTGTATACTTGT	282
Qy	185	acctggttcacatgacccctgtaaccaaacacgcgctacccgtgtgtgaattgggaat	244
Dd	283	ACCTGGTTCCTATGTCACCTGTAAACAAAGAACCGGCTCACTGTGTGGATTTGGGGAAT	342
Qy	245	tcaaacccctctggaacatctggccctggagcttgggaagcttgaacattacaagta	304
Dd	343	TCAAACTCTGTGGACATCTTGGCCCTGAGACTTGGGAAGTTGAACATTACGATATCA	402
Qy	305	gagctctcaaaaaacaacatccaaggaactatacccttcgacttggaaatctgaaagt	364
Dd	403	GAGCTCTTCAAAAACAAACATCCAAAGGAACATACCTTCCAACTGGAAATCGAAGAAAT	462
Qy	365	ctcacacagcttggagatctgtacaacaacaactcttaagagaagttccaccctcttggga	424
Dd	463	CTCATCACTTGGATCTGTCTGTACAAACAACTTACAGGGAATGTCCACTTCTTTGGGA	522
Qy	425	aaattgaagctctggtcttctttagcggttaatgaaacaggaattgacggggccaatccct	484
Dd	523	AAATTGAAGTCTGTGCTTTTTCAGGCTTAATGACAAACGATTGAGCGGGCCAATCCCT	582
Qy	485	agagacacacactgacatcccaagccttaagttgtatgtctccaagcaatgaatttgyt	544
Dd	583	AGAGACATCTACTGCACATCCCAAGCCTTAAGTTGTATGTATGTCTCAACGATGATTTGTGT	642
Qy	545	ggacaacatcccaacaacaacggaacctttgtctcaatctcccttcagaacatttggaaacaac	604
Dd	643	GGAAACAATCCCAACAAAGGACCTTTGTCTACATTCCTTTACGAACCTTTGGAACAAC	702
Qy	605	ccgaggttgggaaggacgggaattactcggtcttcaagctcagcaactaaactgacactga	664
Dd	703	CCGAGTGTGGAGGACCGGATTTACTGSGTCTTGCAAGTCACGACATTAACCTGCACTGA	762
Qy	665	aaaaatctggcaaacactgynaatgaaagaattggggggtgacctgttaagaacacttcacc	724
Dd	763	AAAAATTTGGCAAAACCTTAATAATGAAGATTTGGGGGGAGACCTTGTAAAGAACTTCACC	822
Qy	725	accttatcaaatatcacatctcactatgtaataagtatataatgttagtccaaaaaanaaa	784
Dd	823	ACTTATTCAAATATATACATCTATCTATGTATAAAGTATATATATATGTAGTCCAAAAAANA	882
Qy	785	aaaa 788	
Dd	883	TGAA 886	

RESULT	3	
ID	V06587	standard; cDNA to mRNA; 981 BP.
AC	V06587;	
DT	03-AUG-1998	(first entry)
DE	Arabidopsis thaliana SERK LRR homologous EST clone.	
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos	
KW	plant breeding; leucine-rich repeat; ss.	
OS	Arabidopsis thaliana.	
FH	Key	Location/Qualifiers
FT	CDS	104..760
FT		/tag= a
FT		/note= "shows high homology to SERK"
PN	W09743427-A1.	
PD	20-NOV-1997.	
PF	13-MAY-1997; E02443.	
PR	14-MAY-1996; GB-010044.	
PA	(NOVS.) NOVARTIS AG.	
PI	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;	
DR	WPI; 98-086529/08.	
DR	P-PSDB: W47019.	
PT	Production of apomictic seeds - useful in plant breeding	
PS	Claim 28; Pages 75-77; 123pp. English.	
CC	The sequence is that of an EST clone showing high homology to	
CC	SERK LRR (leucine-rich repeat) sequences.	

Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 94.0%; Score 742; DB 1; Length 981;
Best Local Similarity 96.8%; Pred. No. 26-214;
Matches 757; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 5 cccacgctccggaacatacgttgagagctctcgagcttcgttaactaactta 64
DB 101 CACATGCGCTCGAAACATATCGGTGGACCTTCGCAAGCTTCGTAACCTTA 160
OY 65 gctttgattacactgctgacgaagaactccgaaggaagatgcttcaactcttcg 124
DB 161 GCTTGTATCACTCGTGTGAGCAAGCAACTCCGAGGAGATCTTCAACCTTCG 220
OY 125 agtttaacagatccggaacatgcttcacagagcttggaatccaactctgt 184
DB 221 AGTTGACAGATCCAGACCATGCTCCAGAGCTGGGATCCAACTCTGTATCCT 280
OY 185 accgtgtccatgctacactgttaacaaagaaacggctacactgcttgattgg 244
DB 281 ACCTGTTCCATGTCACCTGTAAACCAAGAACCGGCTCCTGTGTGATTTGG 340
OY 245 tcaaacctctctgacatacttcgctgagcttggaagcttgaacattacagata 304
DB 341 TCAAACTCTCTGACATCTTGCGCTGAGCTTGGAAGCTTGAACATTTACGTA 400
OY 305 gagctctacaacaaacacatccaaaggaactataccttcggaacttgaaat 364
DB 401 GAGCTTCAACAAAACAAATCCCAAGGAATATCTCCAACTTGGAAATCGAAG 460
OY 365 ctcatcagcttgatctgtacaacaacatcttacaaggaagatctccactctt 424
DB 461 CTCTACAGCTTGATCTGTACAAACAATCTTACAGGATTAATCTCCACTT 520
OY 425 aaattgaagctctgctcttcaagccttaagaaacccgattgacggggccaa 484
DB 521 AAATTGAAGTCTCGTCTTTTACGGCTTAATGACAAACGATGACCGGT 580
OY 485 agagcaactcagctgaatcccaagccttaagcttgatgttccaaagcaatgt 544
DB 581 AGAGCACTCAAGCGAATCCAAAGCTTAAGTTGTACCTTCACATGATTTGT 640
OY 545 ggaacaatcccaacaaacaggaactctgtcacatcttcaagaacttgaa 604
DB 641 GGAACAATCCCAAAACGGAACCTTGTCTACATCTTACAGAACTTGAGAAC 700
OY 605 ccgaggtctgagggagccggaattactcgtgtcttgcaagctacgaacta 664
DB 701 CCGAGATTGGAGGAGCCGGAATTAAGTGTGCAAGCTACGACACTAACCT 760
OY 665 aaaaattgcaaaaccttgaanaaataagaatgggggtgaccttgaagaac 724
DB 761 AACAACTGGCAAAACCTGAAATGAAGATGGGGGTGACCTTGAAAGAACT 820
OY 725 actttataaatatcacactctacactgtataaagatatatagtgtccaaaa 784
DB 821 ACTTATCAAAATACATCTATTATGTAATATATATATGTAATAAACA 880
OY 785 aa 786
DB 881 AA 882

RESULT 4

ID V06586 standard; cDNA to mRNA; 1106 BP.

AC V06586;

DT 03-AUG-1998 (first entry)

DE Arabidopsis thaliana SERK LRR homologous EST clone.

KW receptor kinase; apomixis; apomictic; seeds; production; embryos;

KW plant breeding; leucine-rich repeat; ss.

OS Arabidopsis thaliana.

Key Location/Qualifiers
FT 142..798
FT /tag=a
CDS /note="shows high homology to SERK"
MO9743427-A1.
FT 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PI (NOVS) NOVARTIS AG.
PA De Vries SC, Hecht VFG, Schmidt BDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: M47018.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 26; Pages 71-73; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;

Query Match 90.5%; Score 714.2; DB 1; Length 1106;
Best Local Similarity 96.2%; Pred. No. 5.2e-206;
Matches 753; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

OY 5 cccacgctccggaacatacgttgagagctctcgagcttcgttaactaactta 64
DB 139 CACATGCGCTCGAAACATATCGGTGGACCTTCGCAAGCTTGAACCTTA 198
OY 65 gctttgattacactgctgacgaagaactccgaaggaagatgcttcaactcttcg 124
DB 199 GCTTGTATCACTCGTGTGAGCAAGCAACTCCGAGGAGATCTTCAACCTTCG 258
OY 125 agtttaacagatccggaacatgcttcacagagcttggaatccaactctgt 184
DB 259 AGTTGACAGATCCAGACCATGCTTCCAGAGCTGGGATCCAACTCTGTATCCT 318
OY 185 accgtgtccatgctacactgttcaacaaacccgctacactgcttgattgg 244
DB 319 ACCGTTCCATGTCACCTGTAAACCAAGAACCGGCTCCTGTGTGATTTGG 378
OY 245 tcaaacctctctgacatacttcgctgagcttggaagcttgaacattacagata 304
DB 379 TCAAACTCTCTGACATCTTGCGCTGAGCTTGGAAGCTTGAACATTTACGTA 438
OY 305 gagctctacaacaaacacatccaaaggaactataccttcggaacttgaa 364
DB 439 GAGCTTCAACAAAACAAATCCCAAGGAATATCTCCAACTTGGAAATCGAAG 498
OY 365 ctcatcagcttgatctgtacaacaacatcttaaggaagatgttccactctt 424
DB 499 CTCTACAGCTTGATCTGTACAAACAACCTTAAGGATGATGCCACTTGTGGA 558
OY 425 aaattgaagctctgctcttcaagccttaagaaacccgattgacggggccaa 484
DB 559 AAATTGAAGTCTCGTCTTTTACGGCTTAATGACAAACGATGACCGGTCAAT 617
OY 485 agagcaactcactggaatcccaagcctttaaagtgtgtgactcaagaaagatt 543
DB 618 AGACACTCAAGGGAAATCCCAAGCCTTTAAAGTGTGAGCTCTCAAGCAAGATT 677
OY 544 tgaacaatcccaacaaacggaactcttgctacatcttcttaagaacttgaa 603
DB 678 TGGGCAATCCCAACAAAGGAGCCCTTGTGTCATCTTACAGAACTTGAGAAC 737
OY 604 ccgaggtctgagggagccggaattactcgtcttgcaagctacgacataactga 663
DB 738 CCCGAGATTGGAGGAGCCGGAATTAAGTGTGCAAGCTACGACACTTAAGT 797
OY 664 aaaaattgcaaaaccttgaanaaataagaatgggggtgaccttgaagaac 723
DB 798 AACAACTGGCAAAACCTGAAATGAAGATGGGGGTGACCTTGTAAAGCACTT 857
OY 724 caattatcaaatatcacactctatgtataaagatatatagttagtccaaaaaa 783

ID	Accession	Gene	Location/Qualifiers
Db	858	CACCTTATCAAAATACACATCTATATGTAATGAATAAGTATATATAGTATGTAATGAAAAA	917
Qy	784	aaa 786	
Db	918	AAA 920	
RESULT	5		
ID	V06589	standard; CDNA to mRNA; 894 BP.	
AC	V06589;		
DT	03-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK LRR homologous EST clone.		
KW	receptor kinases; apomixis; apomictic; seeds; production; embryos;		
KW	plant breeding; leucine-rich repeat; ss.		
OS	Arabidopsis thaliana.		
FT	key	Location/Qualifiers	
FT	CDS	1..678	
FT	CD	/*tag= a	
FT	CD	/note= "shows high homology to SERK"	
FN	MO9743427-A1.		
PD	20-NOV-1997.		
PF	13-MAY-1997; E02443.		
PR	14-MAY-1996; GB-010044.		
PA	(NOVS) NOVARTIS AG.		
PI	De Vries S, Hecht VFG, Schmidt EDL, Van Holst GJ;		
DR	WPI: 98-086529/08.		
DR	P-PSDB: M47021.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 28: Pages 83-84; 123p; English.		
CC	The sequence is that of an EST clone showing high homology to		
CC	SERK LRR (leucine-rich repeat) sequences.		
SQ	Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;		
Query Match	45.8%; Score 361; DB 1; Length 894;		
Best Local Similarity	72.8%; Pred. No. 1,8e-99;		
Matches	538; Conservative 0; Mismatches 110; Indels 91; Gaps		
Qy	91	ctccgaaggaagatgctctcttaagcgtcttcgcgcgaggttaacaagatccgcgaactgtctt	150
Db	15	CTCCGGAAGGGAGCGCTCTTCACGCCGCTTGCGCGAGCTTATCAGATCCAGCAATGTTGT	74
Qy	151	ccaaagctggatccaaactctgttaactctgttaactctgttaactctgttaactctgttaac	210
Db	75	TCAGAGTTGGGATCAACCTCTGTATATCTGTACTGTGTTCACTGATCATTGTATATCA	134
Qy	211	agacaacgcgcacactcgttgttgatttggtgggaattcaaacactctctgacactcttgcc	270
Db	135	ACACCAATCAAGTCACTCGTCTGGATTGGGGAATTCAAACTATCTGACATCTAGTAC	194
Qy	271	tgaagctgggaagcttgaacattcaagtatct-----	304
Db	195	TGAAGTGGGGAAGCTTGAAACATTTACAAATATCTGATGAAATCATCACTCTTTGCC	254
Qy	304	-----	304
Db	255	TGATTATCTGAAAACATTACATTATCACTGACACACATATACATTCTTGTGAGTATA	314
Qy	304	---agagctcaaaaaaacaacatccaaggaactataccttcggaacttggaaacttgaa	360
Db	315	TAGTGAACTCTCAAAAAACGAGATTCAGAGAACTATACCTTGTGAGTTGGAAATCTGAA	374
Qy	361	gaatctcaagcttgatctgtgaacaaacaaacttaagggaagcttccacttctt	420
Db	375	GAGTCTAATCTAGTTTGGATCTGTACACACACAACTCACCGGGAAATCCATCTTCTT	434
Qy	421	gggaaaatctgaagctctgtgtcttttaagcgttcaatgaacaacgattgaagggccaat	480
Db	435	GGGAAATATGAAGTCACTGTTTGTGTTGGGGCTTAACGAAACGATTTGACCGGATCTAT	494
Qy	481	ccctagagacactcgtcaatcccaagccttaagttgttgttgtctcaagcaatgatt	540

Db	Accession	Gene	Length	Score	DB 1	DB 2	DB 3	DB 4	DB 5	DB 6	DB 7	DB 8	DB 9	DB 10	DB 11	DB 12	DB 13	DB 14	DB 15	DB 16	DB 17	DB 18	DB 19	DB 20	DB 21	DB 22	DB 23	DB 24	DB 25	DB 26	DB 27	DB 28	DB 29	DB 30	DB 31	DB 32	DB 33	DB 34	DB 35	DB 36	DB 37	DB 38	DB 39	DB 40	DB 41	DB 42	DB 43	DB 44	DB 45	DB 46	DB 47	DB 48	DB 49	DB 50	DB 51	DB 52	DB 53	DB 54	DB 55	DB 56	DB 57	DB 58	DB 59	DB 60	DB 61	DB 62	DB 63	DB 64	DB 65	DB 66	DB 67	DB 68	DB 69	DB 70	DB 71	DB 72	DB 73	DB 74	DB 75	DB 76	DB 77	DB 78	DB 79	DB 80	DB 81	DB 82	DB 83	DB 84	DB 85	DB 86	DB 87	DB 88	DB 89	DB 90	DB 91	DB 92	DB 93	DB 94	DB 95	DB 96	DB 97	DB 98	DB 99	DB 100	DB 101	DB 102	DB 103	DB 104	DB 105	DB 106	DB 107	DB 108	DB 109	DB 110	DB 111	DB 112	DB 113	DB 114	DB 115	DB 116	DB 117	DB 118	DB 119	DB 120	DB 121	DB 122	DB 123	DB 124	DB 125	DB 126	DB 127	DB 128	DB 129	DB 130	DB 131	DB 132	DB 133	DB 134	DB 135	DB 136	DB 137	DB 138	DB 139	DB 140	DB 141	DB 142	DB 143	DB 144	DB 145	DB 146	DB 147	DB 148	DB 149	DB 150	DB 151	DB 152	DB 153	DB 154	DB 155	DB 156	DB 157	DB 158	DB 159	DB 160	DB 161	DB 162	DB 163	DB 164	DB 165	DB 166	DB 167	DB 168	DB 169	DB 170	DB 171	DB 172	DB 173	DB 174	DB 175	DB 176	DB 177	DB 178	DB 179	DB 180	DB 181	DB 182	DB 183	DB 184	DB 185	DB 186	DB 187	DB 188	DB 189	DB 190	DB 191	DB 192	DB 193	DB 194	DB 195	DB 196	DB 197	DB 198	DB 199	DB 200	DB 201	DB 202	DB 203	DB 204	DB 205	DB 206	DB 207	DB 208	DB 209	DB 210	DB 211	DB 212	DB 213	DB 214	DB 215	DB 216	DB 217	DB 218	DB 219	DB 220	DB 221	DB 222	DB 223	DB 224	DB 225	DB 226	DB 227	DB 228	DB 229	DB 230	DB 231	DB 232	DB 233	DB 234	DB 235	DB 236	DB 237	DB 238	DB 239	DB 240	DB 241	DB 242	DB 243	DB 244	DB 245	DB 246	DB 247	DB 248	DB 249	DB 250	DB 251	DB 252	DB 253	DB 254	DB 255	DB 256	DB 257	DB 258	DB 259	DB 260	DB 261	DB 262	DB 263	DB 264	DB 265	DB 266	DB 267	DB 268	DB 269	DB 270	DB 271	DB 272	DB 273	DB 274	DB 275	DB 276	DB 277	DB 278	DB 279	DB 280	DB 281	DB 282	DB 283	DB 284	DB 285	DB 286	DB 287	DB 288	DB 289	DB 290	DB 291	DB 292	DB 293	DB 294	DB 295	DB 296	DB 297	DB 298	DB 299	DB 300	DB 301	DB 302	DB 303	DB 304	DB 305	DB 306	DB 307	DB 308	DB 309	DB 310	DB 311	DB 312	DB 313	DB 314	DB 315	DB 316	DB 317	DB 318	DB 319	DB 320	DB 321	DB 322	DB 323	DB 324	DB 325	DB 326	DB 327	DB 328	DB 329	DB 330	DB 331	DB 332	DB 333	DB 334	DB 335	DB 336	DB 337	DB 338	DB 339	DB 340	DB 341	DB 342	DB 343	DB 344	DB 345	DB 346	DB 347	DB 348	DB 349	DB 350	DB 351	DB 352	DB 353	DB 354	DB 355	DB 356	DB 357	DB 358	DB 359	DB 360	DB 361	DB 362	DB 363	DB 364	DB 365	DB 366	DB 367	DB 368	DB 369	DB 370	DB 371	DB 372	DB 373	DB 374	DB 375	DB 376	DB 377	DB 378	DB 379	DB 380	DB 381	DB 382	DB 383	DB 384	DB 385	DB 386	DB 387	DB 388	DB 389	DB 390	DB 391	DB 392	DB 393	DB 394	DB 395	DB 396	DB 397	DB 398	DB 399	DB 400	DB 401	DB 402	DB 403	DB 404	DB 405	DB 406	DB 407	DB 408	DB 409	DB 410	DB 411	DB 412	DB 413	DB 414	DB 415	DB 416
----	-----------	------	--------	-------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Query Match	9.6%	Score 75.4	DB 1	Length 3176
Best Local Similarity	52.0%	Pred. No. 5e-13		
Matches 169	Conservative	0	Mismatches 156	Indels
				Gaps
				0
QY 231	tgagatttgggggaattcaaacctctctctgagacatcttcgcgcctagcttggaaagctgaac	290		
DB 985	TGTATTTCACAGAGTAACAAGCTGAGCTGGTTCAATTCACACCTGAGCTTGGAAACATGTCA	1044		
QY 291	attacagtatcttagagctctacaaaacaacatccaaaggaactatctactccgaacttg	350		
DB 1045	AACCTCCATTACCTGGAGACTCAATGATATCATCTCACGGGTCAATATACCCACGAGCTTG	1104		
QY 351	gaactctgaagatctctcctcaagcttgagatctgtacacacacatcttaacaggaatgctc	410		
DB 1105	GGAACCTTACGTACACTGTTGTTGATCTGAAATGTGGCCACAAATGATCTGGAAGAGACTATAC	1164		
QY 411	ccactctcttgggaaatattgaagctctgctcttcttaccgctatgatgaacagcatga	470		
DB 1165	CTGATCATCTGTAGCTCTTGCCACAAATCTTMAACGCTTAATATGTTCACTTGGAACAAGTTTA	1224		
QY 471	cggggccaactccctctagagactcaactgcatacccgaagcttaagttgtgtctctca	530		
DB 1225	GTCGACATCTTACCCCGAGGATTTCAAAGCTTGAAGATGATGACTTAACTTAATCTGTCCA	1284		
QY 531	gcaatgatcttgtgtggaacatccc	555		
DB 1285	GCAACATATTCAAAGTGCCATCC	1309		
RESULT 9				
ID X23526				
AC X23526	standard; DNA; 5940 BP.			
DE 17-JUN-1999	(first entry)			
DR O. longistaminta	xa21 gene family member A2 DNA.			
KM xa21: receptor kinase-like protein; multigene family; RIK; rice.				
KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.				
OS Oryza longistaminta.				
PN WO909151-A2.				
PD 25-FEB-1999.				
PF 17-JUL-1998: U14841.				
PR 13-AUG-1997: US-910386.				
PA (RESC) UNIV CALIFORNIA.				
PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,				
PI Wang G;				
DR WPI: 99-204431/17.				
PT New RIK polynucleotides and nucleic acid constructs - used for				
CC generating transgenic plants resistant to Xanthomonas				
PS Claim 1, Page 52-53, 67pp; English.				
CC This invention describes a method for conferring disease resistance in				
CC plants. The invention describes the use of novel genes and proteins				
CC belonging to the Oryza longistaminta and Oryza sativa receptor				
CC kinase-like protein (RIK) xa21 multigene family. Such genes from				
CC cassava, maize and tomato are also described. The genes and proteins can				
CC be used for enhancing resistance to Xanthomonas in a plant, preferably in				
CC rice or tomato.				
SQ Sequence 5940 BP; 1570 A; 1200 C; 1188 G; 1982 T;				

Query Match	9.5%	Score 75:	DB 1:	Length 5940:
Best Local Similarity	52.4%	Pred. No. 8	8e-13:	
Matches 155;	Conservative	0;	Mismatches 150;	Indels 0; Gaps 0
Qy	245	tcaacctctcttgacactcttgcgctgagcttggaagcttgaacattacagtatcta	304	
Db	2406	TCGACACTGTCGGGATCATCTCGCCCTCCCTGGGCAACCTGTCCTTCAGACGCTG	2465	
Qy	305	gagctctcaaaaagaacatcagcagagacataccttcggaacttggaaatctgaagaat	364	
Db	2466	GACCTCAAGCACAACCCCTGTCGGGCAAAATACCTCAGAACTCAGCACTTCAGAGG	2555	
Qy	365	ctcatcagcttggatctgtgtcaacaacaacattcaaggatagtgtccactcttggga	424	

Db 2556 CTCACAACTGACGTACTGATTAATTTCACACAGCCTATCCGGGTAGANTTCCAGCTGCTTTGGGC 2555

Qy 425 aaatgaagctctcgtgctcttttaacggttaagacaacgagattgaaggccaatccct 484

Db 2586 AATCTAACCACTGCTCTCGCTTCTTGAGCTGACTATACACTGCTGTGGAGCATCCCT 2645

Qy 485 agagactcactgcatgcccaagccttaagtgttgatgtctcaagaagaattgtgt 544

Db 2646 TCATCTGGGGCAACTACACAGCCTCCTACTGATCTTGCACCTGGCTGAAATAATGCTGCT 2705

Qy 545 ggaacatcccaaca 559

Db 2706 AGTTCAATCCCTTCA 2720

RESULT 10

T06307

ID T06307 standard; cDNA; 3573 BP.

AC T06307:

DT 14-APR-1996 (first entry)

DE Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.

KW Pathogen resistance; Cf-2.2; tomato; C.fulvum; Avr 4; Avr 9; fungal;

KM leaf mould; variegation; ss.

OS Lycopersicon esculentum.

PN W09531564-A2.

PN 23-NOV-1995.

PF 11-MAY-1995; G01075.

PR 11-MAY-1994; GB-009394.

PR 23-DEC-1994; WO-602812.

PR 31-MAR-1995; GB-006658.

PR 07-APR-1995; GB-007232.

PA (GATS-) GATSBY CHARITABLE FOUND.

PI Hammond-Kosack KE, Jones DA, Jones JDG:

DR WPI: 96-010949/01.

DR P-PSDB: R85299.

PT Increasing plant pathogen resistance by induction of variegation -

PS may lead to acquired resistance to a broad range of pathogens.

PS Claim 9: Page 85-87; 131pp: English.

CC T06307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone.

CC In a new method this gene is expressed highly in genetic constructs

CC which may be used to impart a broad range of pathogen resistance, by

CC induction of variegation, to transgenic plants (or parts or propagules

CC of plants) containing such constructs. Cf-2.2 imparts resistance to the

CC disease caused by the leaf mould fungal pathogen Gladosporium fulvum.

CC C.fulvum contains avirulence (Avr) genes that confer recognition by

CC plants containing Cf-genes, leading to the activation of host defence

CC mechanisms to attack the disease.

CC Sequence 3573 BP: 1032 A; 654 C; 664 G; 1223 T;

	Query Match	Similarity	Score	DB	Length
	Best Local	50.4%	Pred. No. 8	5-12	
	Matches 174	Conservative	0	Mismatches 177	Indels 0
					Gaps 0
Qy	215	aaccgcgcgaactcgtgtagattgysgaattcaaacctctctggacatcttgcgectgag	274		
Db	1787	AACAACTTGCTATCTGTATCTTATCCAAATATACACTTCTGGCCTCTTCTCGAAGA	1846		
Qy	275	cttgcggaagcttgaacattacagatctacagctctacaaaaacacatccaaggaat	334		
Db	1847	ATAGGTTACTGAGTCTCTCTACTTACTATATCTTTGGGAATATACCTCTTATGAGACTT	1906		
Qy	335	ataccttcgcgaacttggaaatctgaaagatctcatcaagcttggattctgtacaaacaat	394		
Db	1907	ATTCCGCTTCATTTGGCAATATGAAAAATCTCGAAGCTCTGATTTTCATGTATTAACAT	1966		
Qy	395	cttacaaggaatagttcccaactctcttggysaaattgaagtccttctttttagagcgtt	454		
Db	1967	CTCATTTGGGGAATTTCTTCATCTGTGTGGAATTTGACATCACTGGAAGTGTATATG	2026		
Qy	455	aatgacaacacgattagcaggggccaatcccttagagcaactcatctgcaatccaagccttaa	514		
Db	2027	CCGAGAAACATTTTGAAGGAAAAGTCCGCAATGTTTGGGTAATATACGTATACCTTCAG	2086		

RESULT	15
ID	V14519
AC	V14519 standard: DNA; 3979 BP.
DT	20-MAY-1998 (first entry)
DE	CF-5 pathogen resistance gene variant #2.
KW	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
OS	tomato leaf mould; Phytophthora resistance; ss.
FT	Location/Qualifiers
FT	Key
FT	653..3560
FT	/*tag= a
FT	/product= CF-5 pathogen resistance gene
PD	MO9743428-A1.
PD	20-NOV-1997.
PR	08-MAY-1997. G01249.
PR	24-SEP-1996. GB-019924.
PR	09-MAY-1996. GB-009681.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PA	Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR	WPI; 98-008895/01.
DR	P-PSDB; W41310.
PT	Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT	for production of transgenic plants resistant to pathogens e.g.
PS	tomato leaf mould C. fulvum in tomatoes
PS	Claim 6: Fig 1b: 75pp. English.
CC	This sequence is an example of the polynucleotide of the invention, and
CC	is able to confer pathogen resistance on a plant. It is one of two tomato
CC	Cf-5 gene variants, which offer resistance against the pathogen
CC	Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC	the gene into plant cells and regenerating plants from the cells;
CC	asexually or sexually produced offspring can also be subsequently
CC	produced. Expression of the gene in plant cells can confer pathogen

[illegible]

Search completed: June 23, 2000, 22:53:17
Job time: 40593 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comphen Ltd.

OK nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:45:22 ; Search time 274.21 Seconds
(without alignments)
374.013 Million cell updates/sec

Title: US-09-180-798-26

Perfect score: 789
Sequence: 1 tcgaccacagcgcgcgcgaa.....agtccaaaaaaaaaaaaa 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/6_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
7: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63.6	8.1	3921	4	US-08-567-375-3 Sequence 3, Appl 1
2	63.6	8.1	3921	4	US-08-587-680A-3 Sequence 3, Appl 1
3	63.6	8.1	5992	3	US-08-475-891A-3 Sequence 3, Appl 1
4	58	7.4	2075	1	US-08-238-163-3 Sequence 3, Appl 1
5	55.8	7.1	1554	4	US-08-587-680A-24 Sequence 24, Appl 1
6	55.4	7.0	1058	1	US-08-238-163-1 Sequence 1, Appl 1
7	55.4	7.0	6256	3	US-08-475-891A-1 Sequence 1, Appl 1
8	55.4	7.0	6256	4	US-08-587-375-1 Sequence 1, Appl 1
9	55.4	7.0	6256	4	US-08-587-680A-1 Sequence 1, Appl 1
10	53	6.7	5733	3	US-08-473-553A-1 Sequence 1, Appl 1
11	50.8	6.4	831	4	US-08-567-375-15 Sequence 15, Appl 1
12	48.4	6.1	2880	4	US-08-666-271-1 Sequence 4, Appl 1
13	48.4	6.1	3905	4	US-08-666-271-1 Sequence 1, Appl 1
14	41.4	5.2	1116	2	US-08-244-646-14 Sequence 14, Appl 1
15	41.4	5.2	2917	2	US-08-592-936B-20 Sequence 20, Appl 1
16	41.4	5.2	2917	4	US-09-111-573-20 Sequence 20, Appl 1
17	40.6	5.1	792	2	US-08-244-646-16 Sequence 16, Appl 1
18	40.6	5.1	792	2	US-08-442-063A-32 Sequence 32, Appl 1
19	33	4.2	291	2	US-08-442-063A-35 Sequence 35, Appl 1
20	33	4.2	567	2	US-08-442-063A-38 Sequence 38, Appl 1
21	33	4.2	711	2	US-08-442-063A-41 Sequence 41, Appl 1
22	33	4.2	849	2	US-08-442-063A-44 Sequence 44, Appl 1
23	33	4.2	924	2	US-08-442-063A-47 Sequence 47, Appl 1
24	33	4.2	1002	2	US-08-442-063A-26 Sequence 26, Appl 1
25	33	4.2	1026	1	US-08-272-919-1 Sequence 1, Appl 1
26	33	4.2	1026	2	US-08-619-916-1 Sequence 1, Appl 1
27	33	4.2	1026	6	PCT-US95-08542-1 Sequence 1, Appl 1

28	33	4.2	1593	7	US-08-480-346-1 Sequence 1, Appl 1
29	31.2	4.0	2401	2	US-08-243-541-1 Sequence 1, Appl 1
30	31.2	4.0	2401	3	US-08-480-344-1 Sequence 1, Appl 1
31	31.2	4.0	2401	3	US-08-371-377-22 Sequence 22, Appl 1
32	30.6	3.9	3850	3	PCT-US95-10509-1 Sequence 1, Appl 1
33	30.2	3.8	2241	6	US-08-866-757-1 Sequence 1, Appl 1
34	30.2	3.8	4203	3	US-08-446-908-1 Sequence 1, Appl 1
35	30	3.8	1607	2	US-08-231-205A-1 Sequence 1, Appl 1
36	30	3.8	1607	4	US-08-871-161-1 Sequence 1, Appl 1
37	30	3.8	1607	4	US-08-602-264A-1 Sequence 1, Appl 1
38	30	3.8	2088	4	US-09-056-075-1 Sequence 1, Appl 1
39	30	3.8	6243	4	US-08-189-256A-16 Sequence 16, Appl 1
40	29.8	3.8	457	3	US-08-189-256A-11 Sequence 11, Appl 1
41	29.8	3.8	1448	3	US-08-096-947-2 Sequence 2, Appl 1
42	29.8	3.8	2363	1	US-07-919-140B-2 Sequence 2, Appl 1
43	29.8	3.8	2363	6	PCT-US93-06939-2 Sequence 1, Appl 1
44	29.8	3.8	2363	6	PCT-US93-06939-2 Sequence 1, Appl 1
45	29.6	3.8	594	1	US-08-289-458-1 Sequence 1, Appl 1

ALIGNMENTS

RESULT 1
US-08-567-375-3
; Sequence 3, Application US/08567375
; Patent No. 5952485
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/567,375
FILING DATE: 04-DEC-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645
FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastien, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058930
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product= "Xa-21"
US-08-567-375-3

Query Match 8.1%; Score 63.6; DB 4; Length 3921;
Best Local Similarity 50.3%; Pred. No. 7.1e-11;
Matches 156; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 248 aacctctctgacatcttgccgtgagcttgggaagcttgaacattacagatctagag 307
DB 1159 AAGATCACAGGAACATTCGAGAGATTTGGCATCTTATGCTTACMACATCTCTAT 1218
QY 308 ctctcaaaaacacatcccaagagactatcttcggaacttggaaacttggaaactc 367
DB 1219 CTCTGCACACACATTTTCAGAGGCTCTTCATCATCGTTGGCGAGCTTAACCTTA 1278
QY 368 atcagcttgatctgtacaacaacatctacagagatgtcccacttcttgggaaaa 427
DB 1279 GGCATTCTACCGCCCTACGAAACAACTTGAGCGGTTCATCCCGTTGGCCATAGAAAT 1338
QY 428 ttgaagctctggtctttaaagcttaataagaccgatgacggggccacccctaga 487
DB 1339 CTTCCTGAATTAATATCTTACTCTCGCACACAAATTCAGTGGTGGATACCATAC 1398
QY 488 gcactcactgaaatcccaagccttaagttgttgatgtctcaagcaatgattgttga 547
DB 1399 ACACCTCAACCTCACAAACTTGTGTCTATTAGGCTTCAACTATATACCTTAGTGT 1458
QY 548 acaatcccaa 557
DB 1459 CCAATACCA 1468

RESULT 2
US-08-587-680A-3
Sequence 3, Application US/08587680A
Patent No. 5977434

GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,680A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645

FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/567,375
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058940US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 3921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..2676, 3520..3918)
OTHER INFORMATION: /product= "Xa-21"
US-08-587-680A-3

Query Match 8.1%; Score 63.6; DB 4; Length 3921;
Best Local Similarity 50.3%; Pred. No. 7.1e-11;
Matches 156; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 248 aacctctctgacatcttgccgtgagcttgggaagcttgaacattacagatctagag 307
DB 1159 AAGATCACAGGAACATTCGAGAGATTTGGCATCTTATGCTTACAACTCTCTAT 1218
QY 308 ctctcaaaaacacatcccaagagactatcttcggaacttggaaacttggaaactc 367
DB 1219 CTCTGCACACACATTTTCAGAGGCTCTTCATCATCGTTGGCGAGCTTAACCTTA 1278
QY 368 atcagcttgatctgtacaacaacatctacagagatgtcccacttcttgggaaaa 427
DB 1279 GGCATTCTACCGCCCTACGAAACAACTTGAGCGGTTCATCCCGTTGGCCATAGAAAT 1338
QY 428 ttgaagctctggtctttaaagcttaataagaccgatgacggggccacccctaga 487
DB 1339 CTTCCTGAATTAATATCTTACTCTCGCACACAAATTCAGTGGTGGATACCATAC 1398
QY 488 gcactcactgaaatcccaagccttaagttgttgatgtctcaagcaatgattgttga 547
DB 1399 ACACCTCAACCTCACAAACTTGTGTCTATTAGGCTTCAACTATATACCTTAGTGT 1458
QY 548 acaatcccaa 557
DB 1459 CCAATACCA 1468

RESULT 3
US-08-475-891A-3
Sequence 3, Application US/08475891A
Patent No. 5859339

GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:


```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-05891005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5992 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(512..3149, 3993..4393)
; OTHER INFORMATION: /product-"RRK-B"
; OTHER INFORMATION: /note-"Xaz1 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
; OTHER INFORMATION: sativa)"
US-08-475-891A-3

```

```

Query Match      8.1%; Score 63.6; DB 3; Length 5992;
Best Local Similarity 50.3%; Pred. No. 9.1e-11;
Matches 156; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

```

QY 248 aacctctctgacatcttgccctgagctgagctggaagctggaacattacagatcagag 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1670 AAGATCAGAGGAGGATTCGAGGATGATGCGCATCTTATGCGCTTAACAATCTCAT 1729

QY 308 ctctcaaaaacaacatccaggaactatacctccggaacttggaatctgaagaatctc 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1730 CTCTGCAACAACATTCAGAGGCTCTCTCATCTCGTTGGGCGAGCTTAAAACTTA 1789

QY 368 atcagcttgatctgtacacaacaatcttaccagagatagttccactctcttggaaaa 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1790 GGCATTCTACTCGCTTCGAAAACAACACTTGAGCGGCTTCGATCCCTTGCCATAGGAAT 1849

QY 428 ttgaagctctgtctcttttaaggcttaatagacacagattgacggggccaatccctaga 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1850 CTTACTGAACTTATATCTTACTGCTGCGGCAACAATAATTCAGTGTGATACCATAC 1909

QY 488 gcaatcgaatcccaacgcttaagttggtatgctcgaagcaatgatttggtgga 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1910 ACACTTCGAACCTCAAAACTTGTGTCTCATTAGGCTTCAACTTAATTAACCTTAGTGT 1969

QY 548 acaatcccaa 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1970 CCAATACCA 1979

```

```

RESULT 4
US-08-238-163-3
; Sequence 3, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAVITCH, John M.
; APPLICANT: POWELL, Ann
; APPLICANT: STOFZ, Henrik

```

```

; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
; TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend Kourile and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,163
; FILING DATE: 03-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 421..1401
; US-08-238-163-3

```

```

Query Match      7.4%; Score 58; DB 1; Length 2075;
Best Local Similarity 50.5%; Pred. No. 3.3e-09;
Matches 196; Conservative 0; Mismatches 185; Indels 7; Gaps 2;

```

```

QY 171 ttgtatctctgtacctgttcacatgacacgtgaacaaacacacgcgtcacctg 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 576 TTCTGTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 631

QY 231 tggatttgsggaattcaaacctctctgacatcttgagcctgagcttggaagttgaac 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 632 TCACCGCTTTCAGGCAATATCTCGGCAATTCGCGGAGCCGTGAGACCTTCAT 691

QY 291 atttaagatctagaagctctaca--aaacaacatccaaagaaactatcctcgaac 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 692 ATCTCAAAACATTTGATTTTCATGTTACTATCTCACCGGAACAATTCACCTGCACAA 751

QY 348 ttgaaatctgaagatctcagccttgatctgtatcaaaaacatcttaccagagtag 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 752 TTGCGAGCTCACAATTCACAATGTTAAGGCTTAGCTTACCTTACCAAGTCCGA 811

QY 408 ttccactctcttggaataatgaagctctgtctcttttaagcgttaatagaacagat 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 812 TCCCTGATTCCTTATGCTGAGTAAGAAATTTGACGTTGCTGAGTTGATTAACATCAAT 871

QY 468 tgaaggggccaatccctaagacatcgaatcccaagcctaagttgtatgct 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 872 TTACCGGAACATTCCTTCTCCCTCTCAGCTTCGAAATTTGCTAGCATGTACTTAG 931

QY 528 caagcaatgatttggtgaacatcc 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 932 ATGTAACAACTCACCGGAACAATAC 959

```

```

RESULT 5

```

US-08-587-680A-24
; Sequence 24, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA (partial)
; US-08-587-680A-24

Query Match 7.1%; Score 55.8; DB 4; Length 1554;
Best Local Similarity 49.5%; Pred. No. 1.5e-08;
Matches 144; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 250 cctctcggacatcttgcgcctgagcttggaaagcttgaaacattacagatcttaagct 309
DB 18 CTTTCTGCGGACCTTCTAGTGTGCTATTGGAACCTATTTCAGGGCTGAGAGATCTTGTGT 77
QY 310 ctacaacaacacatcccaaggaaactacctccgaacttgcgaacttgaagaatctcat 369
DB 78 AACGGAATGTTCTCAGGTGATATCCCTCTGATATTGGCAGACTAAGAGCATCTT 137
QY 370 cagcttgatctgacaacaacaaacttacaaggatagttcccactcttctggaaat 429
DB 138 AAAGCTGACCTGAGTAGAACAACACTCTCTGCGACAACTCCCTCAGATGTGTAAC 197
QY 430 gaatctcctgctcttcttaagagcttaatgacaacagattgacggggcaatccctagagc 489

DB 198 TCTTCTTACTTACTTGGATTGAGCCCAATCACTTTGTGTCATCCAGTCA 257
QY 490 actcaactgcaatcccaagccttaagttgtatgtatcgaagaatgatt 540
DB 258 AATTGCTCAATTCACATCTTAATTACATCAATATTCTCGAATCACTT 308
RESULT 6
US-08-238-163-1
; Sequence 1, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAVITCH, John M.
; APPLICANT: POWELL, Ann
; APPLICANT: STOTZ, Henrik
; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
; TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Knourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,163
; FILING DATE: 03-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1058 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..1023
; NAME/KEY: misc_feature
; LOCATION: 1..1058
; OTHER INFORMATION: /standard_name= "Pear PGIP cDNA"
US-08-238-163-1

Query Match 7.0%; Score 55.4; DB 1; Length 1058;
Best Local Similarity 51.4%; Pred. No. 1.6e-08;
Matches 128; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 241 gaattcaaacctctctgacatcttgcgcctgagcttggaaagcttgaacattacagta 300
DB 342 GCAACCAATCTCACTGCGCCCAATCCAAACCGGCATTTGCCAAGCTCAAGATCAAGTC 401
QY 301 tctagagctctacaanaaacacacatccaaggaactatactcttcgaacttggaaatctgaa 360
DB 402 TCTGAGCTCAGCTGAGCAACACTCTCAGGCTGTGTCCTGACTCTTCAGCCAACTCAA 461
QY 361 gaatctcctgcttctgtatctgtacaacaacacattcaaggagatagttcccactctt 420

Db 462 GAACTTACATTCCTCACCCTCTTCACACACCTCACCAGGCGCATCCAGCTCGCT 521
Oy 421 gggaattgaagctcgtcttcttaagccttaagcaacagcttgaagggccaat 480
Db 522 TTCTGACTCCCAAACTCGCGCTCTTCGTCTAGACCCGATTAAGCTCAAGTCAAT 581
Oy 481 ccctagagc 489
Db 582 TCCGATATC 590

RESULT 7

US-08-475-891A-1
; Sequence 1, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
; OTHER INFORMATION: sativa)"
US-08-475-891A-1

Query Match 7.08; Score 55.4; DB 3; Length 6256;
Best Local Similarity 50.28; Pred. No. 4.5e-08;
Matches 165; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

Oy 231 tggattgggaattcaacctctcgtgacatcttgagccttgagccttgaggaagctgaac 290
Db 2735 TGGACTTGGAGAAATTAACCTGGGGGAGTCTCTTAATTCGTTTCCAAATCTTCCA 2794

Oy 291 att---taccgtatctagaagctctcaaaaacacatccaaaggaactatacttcggaac 347
Db 2795 CTTCGGCTAGATTCTCTTGACCTTGATTTGAATTAATATACAGGAAGCATTTCCAAAGATA 2854
Oy 348 ttggaattcgaagaatctcaatcgaagcttgagctctgatacaacaacaattcaagggatag 407
Db 2855 TTGGCAATCTTAATGGCTTACAAATCTTAATCTCTGCAACAAATTTCCAGAGGCTCAC 2914
Oy 408 ttccactcttgaggaaattgaagctctgctcttcttaagccttaatgacaacgat 467
Db 2915 TTCCATCATTCGTTGGGCGAGCTTAGAAACTTAGGCATTCGTGCGCCACGAAACAACT 2974
Oy 468 tgaaggggccaatccctagagcactcaactgcaatcccaagccttaagttgagtct 527
Db 2975 TGAAGGCTTGATCCATTCGCGCATAGGAATCTTAATCTTAATCTTAATCTTAATCTGCTCG 3034
Oy 528 caagcaatgattgtgtggaacatccca 556
Db 3035 GCACCAACAAATTCAGTGGTTGATACCA 3063

RESULT 8

US-08-567-375-1
; Sequence 1, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)


```

; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/REF/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277295
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
;
; US-08-473-553A-1
;
Query Match
Best Local Similarity 6.7%; Score 53; DB 3; Length 5733;
Matches 128; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
;
QY 310 ctcaaaacaacatcccaaggaacatcatccttcgaacttgaaatcgaagaatctcat 365
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3120 ctctcctcaacacagctacacccgggtggttccacgcagcttccgtgttaacaaagcttga 3179
;
QY 370 cagcttgagatctgtacaacaacatctacaggaatagttccactctcttgaggaaat 429
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3180 gatccctcgacatggcagctgtacactcacccgagacagatccgagatttaagtaacct 3239
;
QY 430 gaagctctggtctttttagcgcttaagaaacacgcatgacggggccaatccctagaagc 489
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3240 gaaacatctacatctctttttcttcacacacacattaacccgctcatatcacccgga 3299
;
QY 490 actactgtaaccccaagccttaagttgtagtctcaagcaatgattgtgtggaac 549
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3300 gctttccggtttagtcagcttgaattcgtcatattatcaatcaatcgttaacccgga 3359
;
QY 550 aatcccaacaac 562
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3360 aatccctcaaac 3372
;
;
RESULT 11
; US-08-567-375-15
; Sequence 15, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants

```

```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; US-08-567-375-15
;
Query Match
Best Local Similarity 6.4%; Score 50.8; DB 4; Length 831;
Matches 139; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
;
QY 255 ctgagacatcttgagcctgagcttgaggagcttgaacattacagttatctagagcttaca 314
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 33 ctggtgcacttctttagtcttcttttgaaactatctcagggctgaagaaatctgttactg 92
;
QY 315 aaaaacaacatcccaaggaactatacttcgaacttggaaatctgaagaatctcatcagct 374
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 93 gaaatggttttcagaggtatattccctctgatttgacagactaaagacattttaaac 152
;
QY 375 tggactctgaacaacacatctacaggaatagttccactctcttgaggaaatgaagt 434
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 tggactcagtgaaacacattctctgacaaatccctctcagattggaactgttctt 212
;
QY 435 ctctgtctttttagcgcttaatgacaacgattgacggggccaatccctagaagcacta 494
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 ccttaacttacttgatttgagccaaatcaattcttgcttgatccaggttcaattg 272
;
QY 495 ctgcaatcccaagccttaagttgtagtctcaagcaatgattt 540
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 273 ctcaattcacttcttaattatcaatatttcttggaatcactt 318
;
;
RESULT 12
; US-08-666-271-4
; Sequence 4, Application US/08666271
; Patent No. 5920000
; GENERAL INFORMATION:

```


Db 2973 TGCACCTGACGTTGCAATGTCGACATGTCGAAGTCATATACCGCATCAT 3032
Qy 421 gggaaattgaagtccttgccttttaagcgtaatagcacccgattgacggggccaat 480
Db 3033 TCAAAATTTATCAGTATCGAACTCTTGATCTTCATCTAATAAATACGCGAATA 3092
Qy 481 ccctagagcactcactgcaatcccaagccttaagtgtgctcacaagcaatgatt 540
Db 3093 TCCGACAGCTGTCATCCCTCACATTCCTTGAACTCTAATAATCTCTCACAAATCATCT 3152
Qy 541 gtgtggacaatcccaacaagcgaccttgcct 574
Db 3153 TGTGTGATGATCCCAAGAAACAAATTTGAT 3186

RESULT 14
US-08-244-646-14
; Sequence 14, Application US/08244646
; Patent No. 5744692
; GENERAL INFORMATION:
; APPLICANT: Cervone, Felice
; APPLICANT: De Lorenzo, Giulia
; APPLICANT: Salvati, Giovanni
; APPLICANT: Albersheim, Peter
; APPLICANT: Darvill, Alan
; APPLICANT: Bergmann, Carl
; TITLE OF INVENTION: Nucleotide Sequences Coding An
; TITLE OF INVENTION: Endopolygalacturonase Inhibitor
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sally A. Sullivan
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,646
; FILING DATE: 06-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM 91A 000915
; FILING DATE: 06-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/IT/00158
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sullivan, Sally A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 19-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Phaseolus vulgaris
; STRAIN: Saxa
; IMMEDIATE SOURCE:
; CLONE: lambda pGIP-3.3

; FEATURE: CDS
; NAME/KEY: 1.1026
; LOCATION: 1.1026
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1027..1116
; US-08-244-646-14

Query Match 5.2%; Score 41.4; DB 2; Length 1116;
Best Local Similarity 48.1%; Pred. No. 0.00062;
Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Qy 313 caaaacacatcccaaggaactatacctccgaacttggaaatcgaatcctcag 372
Db 342 CATCAATTAACCTGTCGTCGAATCCCGCCGCAATCCGCTAATCAACCCACCACTCA 401
Qy 373 ctggatctgtacaagaacatctacagggatgtccacactcttgggaaattgaa 432
Db 402 TCTCTATATCAGTATCGAACTCTTGATCTTCATCTAATAAATACGCGAATA 461
Qy 433 gtctcgtcttctttaaaggcttaatagaacagatgacgggacacatccctagagact 492
Db 462 AACCTGTCACCCCTGCACTTCTCTACAAAGCCCTCTCCGACACCCCTCCCTCAT 521
Qy 493 cactgaatcccaagccttaagtgtgatagtctcacaagcaatgattgtggaacat 552
Db 522 CTCTCTCTCCCAACCTCGAGAAATCATTTGACGCGCAACCGAATCTCGGCGCAT 581
Qy 553 CCC 555
Db 582 CCC 584

RESULT 15
US-08-592-936B-20
; Sequence 20, Application US/08592936B
; Patent No. 5783393
; GENERAL INFORMATION:
; APPLICANT: Kelloeg, Jill A.
; APPLICANT: Bestwick, Richard K.
; TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR
; TITLE OF INVENTION: REGULATED EXPRESSION OF TRANSGENES IN PLANTS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,936B
; FILING DATE: 23-JAN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, Susan T.
; REGISTRATION NUMBER: 38,443
; REFERENCE/DOCKET NUMBER: 4257-0012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0980
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:35 ; Search time 6198.48 Seconds

(without alignments)
515.932 Million cell updates/sec

Title: US-09-180-798-26

Perfect score: 789
Sequence: 1 tcgaccacgcgcgcgcgaa.....agtccaaaaaaaaaaaaa 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

Db 180 AACTTGAGAACACCGAGGTTGGAGGACCGGAATTACTCGCTTTCGCAACCTACGAC 239
 OY 650 actactgcaccctgaataaatggcaaacctgaataatgaagaattgggggtgacattg 709
 Db 240 ACTAATCTACCTGCTAATAAATTTGGCAAACTGAAATGAGAAATTGGGGGGGACCTTG 239
 OY 710 taagaacactcaccacttatcaataatcacatctactatgttaataatgatatatgt 769
 Db 300 TAAGAACACTTACACCACTTATCAAAATATCATCTACTATGTAAATAGTATATATGT 359
 OY 770 agtcc 774
 Db 360 AGTCC 364

RESULT 4
 A1100682 348 bp mRNA EST 21-AUG-1998
 LOCUS 33803 lambda-PRL2 Arabidopsis thaliana cDNA clone 185LXP 3', mRNA
 DEFINITION
 ACCESSION A1100682
 VERSION A1100682.1 GI:3449392
 KEYWORDS EST.
 SOURCE
 ORGANISM Arabidopsis thaliana
 thale cress.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsids.
 1 (bases 1 to 348)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
 Retzel,E. and Somerville,C. Methods for accessing results from
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 On May 5, 1995 this sequence version replaced gi:797667.
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313t@msu.edu
 The sequence entry for this EST has been reverse complemented and
 is being submitted in the sense orientation.
 Seq primer: M13-21.

FEATURES
 source
 1..348
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="185LXP"
 /clone_1lb="lambda-PRL2"
 /note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not;
 lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRU's lambda zip-lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA."

BASE COUNT 110 a 74 c 66 g 93 t 5 others
 ORIGIN

Query Match 42.1%; Score 332; DB 42; Length 348;
 Best Local Similarity 99.1%; Pred. No. 3, 2e-81;
 Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 440 gtcttttaaggcttaatgacacccgatgaggggccatccctagacactcaatgca 499
 Db 14 GCTTTTACGCGCTTAATGACACCGAATGACGGGGCCAACTCCNTAAGACATCAGCA 73
 OY 500 atcccaagccttaagtggtgtgctgctcaagcaatgattgtgtggaacatccaca 559
 Db 74 ATCCCAAGCCTTAAGTGTGATGCTCAAGCAATGATTGTGTGCAAAATCCCAACA 133
 OY 560 aacggaccttgcctcacattccctttacagaccttgaagaaacccgaggttgaggga 619
 Db 134 AACGACCTTTGCTCACATTCCTTACAGAACTTGAAGAAACCCGAGGTGAGAGGA 193
 OY 620 ccggaattactcgtctctgcaagctagacactacgccttgaataatggcaaac 679
 Db 194 CCGGAATTAATCTGNTCTGCAAGCTAGCAGCACTACCTGCAAAATGTCGCAAAAC 253
 OY 680 ctgaataagaattgggggtgacctgtgaagaacactcaccactttacaatc 739
 Db 254 CTGAATAATGAATTTGGGGGTGANTCTGTAGAACACTTCACCACTTATCAAAATATC 313
 OY 740 acactctatgttaataatgatatatgtagtc 774
 Db 314 ACATCTACTATGTAAATAGTATATATATAGTATAGTCC 348

RESULT 5
 H37300 353 bp mRNA EST 30-DEC-1997
 LOCUS H37300
 DEFINITION 15429 lambda-PRL2 Arabidopsis thaliana cDNA clone 179J197T, mRNA
 sequence.
 ACCESSION H37300
 VERSION H37300.1 GI:906799
 KEYWORDS EST.
 SOURCE
 ORGANISM Arabidopsis thaliana
 thale cress.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsids.
 1 (bases 1 to 353)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
 Retzel,E. and Somerville,C. Methods for accessing results from
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 On May 8, 1995 this sequence version replaced gi:801146.
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313t@msu.edu
 Seq primer: T7 dye primer.

FEATURES
 source
 1..353
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="179J197"
 /clone_1lb="lambda-PRL2"
 /note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not;
 lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRU's lambda zip-lox. The cDNA

Inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. " 89 t 14 others
BASE COUNT 98 a 85 c 67 g 89 t 14 others
ORIGIN

Query Match 41.5%; Score 327.8; DB 23; Length 353;
Best Local Similarity 94.1%; Pred. No. 4.6e-80;
Matches 332; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 311 tacaaaaaacatccaaagaaactatatactctcgaacttgaaatctgaagaatctcacc 370
Db 1 TACAAAAACATCCAAAGAACTATACCTCCGAATTGGAATCTGAAGANTCTCATC 60
QY 371 agcttgatctgtacaacaaacatcttaaggaatagttccactctcttggaagaattg 430
Db 61 AGCTTGATCTGTACAAACAACTTACAGGATAGTCCACTTTTGGGAAATATG 120
QY 431 aagctctgtctcttttaagcgttaaatgacacacgattgacggggcccaatccctagaag 490
Db 121 ANGTCTGTGCTTTTNAACGCTTAATGACACGATGACCGGTCCATCCTTACAGCA 180
QY 491 ctcaactgaatcccaagccttaagttgtatgctcgaagcaatgattgtgtggaaca 550
Db 181 CTCACGGCAATCCCAAGCTTAAGTGTGANTCTCANGCAATGATTTGTGTGGAACA 240
QY 551 atcccaaaaacagcctcttgctcacatcccttaacagactttgaagaaacccgagg 610
Db 241 ATCCCAACAAAGCAGCCCTTTGCTCACAATCTTTACAGAACTTTGNGAACAACCGGNA 300
QY 611 ttggaaggacccgaattactcgtctgtcgaagctacagacactactgacactg 663
Db 301 TTGGAGGGGCGGANTTACTCGTCTGCAAGANTACGGCACTAATGCAACACTG 353

RESULT 6
AA394359 534 bp mRNA EST 30-OCT-1997
LOCUS 25942 Lambda-PRU2 Arabidopsis thaliana cdna clone 30561T7 3', mRNA
DEFINITION
AA394359
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 534)
REFERENCE
AUTHORS
Newman, T., deBruijn, F. J., Green, P., Keegstra, K., Kende, H.,
McIntosh, L., Ohlrogge, J., Ralke, N., Somerville, S., Thomas, M.,
Retzel, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cdna clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On May 18, 1995 this sequence version replaced gi:811121.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313c@msu.edu
Seq primer: T7.

FEATURES
source
1..534
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="30561T7"

/clone.lib="Lambda-PRU2"
/note="Vector: Lambda Zip-Lox; Site.1: Sal; Site.2: Not;
Lambda PRU2 is a cdna library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark-rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRU's Lambda Zip-Lox. The cdna
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. "

BASE COUNT 151 a 114 c 125 g 126 t 18 others
ORIGIN

Query Match 41.4%; Score 327; DB 33; Length 534;
Best Local Similarity 88.0%; Pred. No. 8.1e-80;
Matches 374; Conservative 0; Mismatches 45; Indels 6; Gaps 2;

QY 369 tcagcttgatctgtacaacaaacatcttaaggaatagttccactctcttggaagaat 428
Db 1 TCAGCTTGATCTGTACAAACAACTTACAGGATAGTCCACTTTTGGGAAAT 60
QY 429 tgaagctctgtctcttttaagcgttaaatgacacacgattgacggggcccaatccctagaag 488
Db 61 TGAAGCTCTGTGCTTTTNAACGCTTAATGACACGATGACCGGTCCATCCTTACAGCA 120
QY 489 caccactgaatcccaagccttaagttgtatgctcgaagcaatgattgtgtgaa 548
Db 121 CACTCAGGCAATCCCAAGCTTAAGTGTGANTCTCANGCAATGATTTGTGTGGAACA 180
QY 549 caatcccaaaaacagcctcttgctcacatcccttaacagactttgaagaaacccgga 608
Db 181 CAATCCCAACAAAGCAGCCCTTTGCTCACAATCTTTACAGAACTTTGNGAACAACCGGA 240
QY 609 ggttggaaggacccgaattactcgtctgtcgaagctacagacactactgacactgaa 668
Db 241 GATTGGAGGAGACCGGAATTCTGCTGCAAGCTTACGCACTAATGCAACACTG 300
QY 669 attgcaaaacccggaatgaatgaattgggggtg---accctgaaagacactcacc 725
Db 301 ACTGCAAAACCTGGAATGAGAGATTTGGGGGGGTGACCTTTGTAAGACACTTCANCA 360
QY 726 cttatacaatatacatctac---tatgtaataagatatatagtgtccaaaaaaa 782
Db 361 CTTTATCCATATTCACATCCATTAANGAATAGTATATAGTGTAAACCA 420
QY 783 aaaaa 787
Db 421 AAAAA 425

RESULT 7
AA443205 608 bp mRNA EST 17-FEB-2000
LOCUS EST308135 tomato mixed elicitor, BTI Lycopersicon esculentum cdna
DEFINITION
AA443205
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 608)
REFERENCE
AUTHORS
Liang, F., Hansen, T. S., Romling, C. M., Craven, M. B., Bowman, C. L.,
Nieman, W., Fraser, C. M., Venter, J. C., Tanksley, S. D.,
Giovannoni, J. J. and Martin, G. B.
Generation of ESTs from tomato callus (mixed elicitor)
Unpublished (1999)

FEATURES
source
1..608
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="30561T7"

COMMENT

On Dec 20, 1995 this sequence version replaced gi:1135853.
 Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@clemson.edu
 Location/Qualifiers
 1..608

FEATURES

source

/organism="Lycopersicon esculentum"
 /cultivar="Rio Grande Flor"
 /db_xref="taxon:4081"
 /clone="CLET43M10"
 /clone_lib="tomato mixed elicitor, BTI"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; CLEF - inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, ETX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."
 BASE COUNT 165 a 126 c 134 g 183 t
 ORIGIN

Query Match 40.3%; Score 318.2; DB 81; Length 608;
 Best Local Similarity 74.4%; Pred. No. 2.2e-77;
 Matches 401; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 80 gtggaagaaactccgaagagatgctcttaacgctcttgccggaggttaacagatcg 139
 Db 69 gtaaggggaattcagaaggagatgcttgcacccctccgacgcttaacgacccg 128
 QY 140 gacatgtctccagagcttgatccaaactctgttaaccttgatctgacatgctc 199
 Db 129 ggtacgctgttacagagctgagatccaaatctgttaaccccttgatctgacatg 188
 QY 200 acctgttaacgaagacacgcgtacatgctgtgtgatttgaggaaattaaacctctg 259
 Db 189 actgtcaacgagatcaatcaagttactgctgctgagatctggaaactcaaaattatctg 248
 QY 260 catctgagctgagctgagagagcttgaaacttaacattacagatcagaagcttaacaaac 319
 Db 249 catttggtactgagcttggaagcttgaacattacagatcagaagcttgaacaaat 308
 QY 320 aacatcgaaggaactatactccgaacttgaaacttgaaacttgaaacttgaaact 379
 Db 309 aatatttcaggaacatcccttaagagctcggttaacgtgaagacccattatgactgat 368
 QY 380 ctgttcaacaaacatctacagagatgattcccaactcttgaggaaattgaagctctg 439
 Db 369 ctgttcaacaaacatctacagagatgattcccaactcttgaggaaattgaagctctg 428
 QY 440 gctctttacagcttaagacacacgattgagcgagccaccccttagagacatcagca 499
 Db 429 gttttcttgcttaaaatgattgagacacacgattgagcgagccaccccttagagac 488
 QY 500 atcccaagccttaaaatgattgagacacacgattgagcgagccaccccttagagac 559
 Db 489 atttttagccttaaaatgattgagacacacgattgagcgagccaccccttagagac 548
 QY 560 aacgagacctttgctcaactctttagacagacatttgagacacacccgaggttgagg 618
 Db 549 tctggtcattttgagcattttctttaaacaatttcgagacacacccgaggttgagg 607

RESULT 8
 AM031110
 LOCUS AM031110 720 bp mRNA EST 15-SEP-1999

DEFINITION

EST274417 tomato callus, TAMU Lycopersicon esculentum cDNA clone
 CLECI3C21 similar to leucine-rich repeat protein (LRR), mRNA
 sequence.

ACCESSION
 VERSION AM031110
 KEYWORDS
 SOURCE EST

ORGANISM

tomato.
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 720)

REFERENCE

Aleale, J., Vredalov, J., White, R., Matern, A.L., Vision, T.,
 Holt, L.E., Liang, F., Upton, V., Craven, M.B., Bowman, C.L., Ahn, S.,
 Romling, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.
 Generation of ESTs from tomato callus tissue
 Unpublished (1999)

TITLE

Generation of ESTs from tomato callus tissue

JOURNAL

Unpublished (1999)

COMMENT

On Dec 20, 1995 this sequence version replaced gi:1135853.
 Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@clemson.edu
 5 prime sequence.

FEATURES

Location/Qualifiers
 1..720
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLECI3C21"
 /clone_lib="tomato callus, TAMU"
 /tissue_type="callus"
 /dev_stage="25-40 days old"
 /lab_host="XLI-Blue MRF"
 /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni laboratory; CLEC - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 200 a 149 c 143 g 228 t
 ORIGIN

Query Match 40.2%; Score 317.2; DB 64; Length 720;
 Best Local Similarity 74.3%; Pred. No. 4.2e-77;
 Matches 400; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 80 gtggaagaaactccgaagagatgctcttaacgctcttgccggaggttaacagatcg 139
 Db 183 gtaaggggaattcagaaggagatgcttgcacccctccgacgcttaacgacccg 242
 QY 140 gacatgtctccagagcttgatccaaactctgttaaccttgatctgacatgctc 199
 Db 129 ggtacgctgttacagagctgagatccaaatctgttaaccccttgatctgacatg 302
 QY 200 acctgttaacgaagacacgcgtacatgctgtgtgatttgaggaaattaaacctctg 259
 Db 189 actgtcaacgagatcaatcaagttactgctgctgagatctggaaactcaaaattatctg 362
 QY 260 catctgagctgagctgagagagcttgaaacttaacattacagatcagaagcttaacaaac 319
 Db 249 catttggtactgagcttggaagcttgaacattacagatcagaagcttgaacaaat 422
 QY 320 aacatcgaaggaactatactccgaacttgaaacttgaaacttgaaacttgaaact 379
 Db 423 aatatttcaggaacatcccttaagagctcggttaacgtgaagacccattatgactgat 482
 QY 380 ctgttcaacaaacatctacagagatgattcccaactcttgaggaaattgaagctctg 439

Db	483	CTGTACAAACAATATTTGGGGACAAATTCCTACTTGTGGAAACCGTGAACAAACCTT	542
Qy	440	gtctttttaagggttaatgacaaacgattgaggggccaattcccttagagactcaetgca	499
Db	543	GTTTCTTGCGTTAATATGATTAACAAGCTAACAGGACCAATCCCAAGAAAGAACTTAATGAC	602
Qy	500	atcccaagccttaagdtgtgatgtctcgaagcaatgatgtgtgtgnaaaatccccaaca	559
Db	603	ATTTCAGCCTGAATCTGTGATGCTCGAATATGATTTGTGTGGAAACAATTCCTACT	662
Qy	560	aaggacacttttctacacttccctttaagaactttgagaacaaccggagtggtggagg	617
Db	663	TCTGTGTCATTTGACATATTCCTCTTAACAATTTGAGCAGCAATTCCTGCACCTTGAAG	720

RESULT	9
LOCUS	A1100679
DEFINITION	A1100679 339 bp mRNA
ACCESSION	33800 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179J11XP 3', mRNA
VERSION	sequence.
KEYWORDS	A1100679
SOURCE	A1100679.1 GI:3449389
ORGANISM	EST.
	thale cress.
	Arabidopsis thaliana

REFERENCE 1 (Pages 1 to 339)
 Neuman T., debuijn, F.J., Green, P., Keegstra, K., Kende, H.,
 Metcosh, L., Ohlrogge, J., Ralthei, N., Somerville, S., Thomasow, M.,
 Ratzel, F. and Somerville, C
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 COMMENT On May 5, 1995 this sequence version replaced g1:797664.

Michigan State University
MSU-OEE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel.: 517-353-0854
Fax: 517-353-9168
Email: 223j3t@clim.c1.msu.edu
The sequence entry for this EST has been reverse complemented and
is being submitted in the sense orientation.
Seq primer: M13-21.

```

location/Qualifiers
1. .339
/organism="Arabidopsis thaliana"
/strain="var columbiana"
/db_xref="taxon:3702"
/clone="17911xp"
/clone_1b="lambda-PRL2"
note="Vector: lambda zip-lox; Site_1: Sal; site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
root roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
Oligo dt primed cDNA."
BASE COUNT      104 a      75 c      66 g      89 t      5 others
ORIGIN

```

Query Match	39.98;	Score 314.6;	DB 42;	Length 339;
Best Local Similarity	97.08;	Pred. No. 2e-76;		

Matches	328;	Conservative	0;	Mismatches	9;	Indels	1;	Gaps	1;
OY	438	tgagccttttaagcgttaatgaacacccgattgaaggggccaa-tccctaaagacatcact							496
Db	2	TGTTCTTTTACGGCTTAATGACAAACCATTAGCGGGCCAAATTCCTTAAGACACACT							61
OY	497	gcacatcccaagccttaagaatgtgtatgtctcaagcaatgaattgtgtgtaaacatccca							556
Db	62	GCATTTCCCAAGCCTTTAAAGTTGTATGATTTCCAGCAATGTTTGTGTGGAAACANTRCCA							121
OY	557	acaaagagaccttgcacatcccttcaagaacttggaaacacccgaggttgag							616
Db	122	ACAACGAGACTTTTCTCACAATTCCTTTACGAACCTTGAAGAACCCGAGGTTGGAG							181
OY	617	ggacccggaattactcgtgtctgcgaagtaagacactaactgacactgaaaaattggcaa							676
Db	182	GGACCGGAAATTAATCTCGGGCTTGCAAGCTACGACACTAATCTGACCTTAAAAATTTGGCAA							241
OY	677	aaccttgaanaatgaagaattcggggggtggaacctgttaagaacacttcacacattcatcaaat							736
Db	242	AANCTGAATGAAGAAATTGGGGGGGGAACCTGTGAAGAACACTTCACCACTTAAAGCAAT							301
OY	737	atcacactcactatgtaataagaatataatagtgtcc							774
Db	302	ANCAACATCTCTATGTAAATPAAGATATATATATGANNCC							339

RESULT	10
LOCUS	AA712221
DEFINITION	AA712221 401 bp mRNA EST 24-DEC-1997 31949 Lambda-PRL2 Arabidopsis thaliana cDNA clone 180C977, mRNA sequence.
ACCESSION	AA712221
VERSION	AA712221.1
KEYWORDS	GI:2722138
SOURCE	EST.
ORGANISM	thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 401) Newman,T., debruin,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlgege,T., Ralshel,N., Somerville,S., Thomasow,M., Ratzel,E. and Somerville,C. Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994) 95148729 On Sep 19, 1997 this sequence version replaced q1:1520373.
TITLE	
JOURNAL	
COMMENT	

```

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lpm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 401
/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone_id="180C97"
/clone_1lb="Lambda-PRL2"
/note="Vector: lambda ZAP-Lox; Site_1: Sal; Site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 2
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark-rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:21:39 ; Search time 12463.9 Seconds
(without alignments)
-69.776 Million cell updates/sec

Title: US-09-180-798-28
Perfect score: 894
Sequence: 1 ggaccgatccaagcctcga.....gtaactggaattcttattta 894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_y1:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_y1:*
32: gb_hcg1:*
33: gb_hcg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hcg3:*
42: gb_hcg4:*
43: gb_hcg5:*
44: gb_hcg6:*

45: gb_hcg7:*
46: em_hcg1:*
47: em_hcg2:*
48: em_hcg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_hcg8:*
53: gb_hcg9:*
54: gb_hcg10:*
55: gb_hcg11:*
56: gb_hcg12:*
57: gb_hcg13:*
58: gb_hcg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	894	100.0	894	5	A67823	A67823 Sequence 28
2	361	40.4	789	5	A67821	A67821 Sequence 26
3	361	40.4	1063	5	A67825	A67825 Sequence 30
4	356.2	39.8	981	5	A67819	A67819 Sequence 24
5	330	36.9	1106	5	A67817	A67817 Sequence 22
6	119.4	13.4	4604	7	LEUREPENE	X95269 L.esculentu
7	106	11.9	2089	5	A67827	U93048 Daucus caro
8	95.4	10.7	1755	8	DCU93048	A67797 Sequence 2
9	93	10.4	936	8	SRU62279	U62279 Sorghum bic
10	79	8.8	4327	32	T5F17	AL049917 Arabidops
11	77.8	8.7	2786	7	AB029327	AB029327 Nicotiana
12	77.8	8.5	142418	7	AP000815	AP000815 Oryza sat
13	73.8	8.3	87434	7	AB010698	AB010698 Arabidops
14	73.4	8.2	3176	5	E12705	E12705 Arabidops
15	73.4	8.2	3176	8	ATU47029	U47029 Arabidops
16	71.6	8.0	4081	5	A67815	A67815 Sequence 20
17	71.6	8.0	98471	8	AC012654	AC012654 Arabidops
18	71.6	7.8	75050	8	AC002334	AC002334 Arabidops
19	70	7.8	100887	50	ATAC011620	ATAC011620 Arabidops
20	69.6	7.7	111945	8	AT1C12	AL022224 Arabidops
21	68.8	7.6	52717	7	AB019227	AB019227 Arabidops
22	68.2	7.6	96475	7	ATF17M5	AL035678 Arabidops
23	68	7.6	81513	50	AC021198	AC021198 Arabidops
24	67.6	7.6	108355	8	AC005957	AC005957 Arabidops
25	67.6	7.5	143186	8	AC005287	AC005287 Arabidops
26	67	7.4	3573	5	A57133	A57133 Sequence 4
27	66.4	7.4	3368	8	U42445	U42445 Lycopersico
28	66.4	7.4	6471	5	A57130	A57130 Sequence 1
29	66.4	7.4	6471	8	U42444	U42444 Lycopersico
30	66.4	7.4	83649	50	AC005395	AC005395 Arabidops
31	66.4	7.4	116944	8	AC008017	AC008017 Arabidops
32	66.2	7.4	99725	8	ATF18F4	AP021637 Arabidops
33	65.8	7.3	2936	8	AF053964	AF053964 Lycopersi
34	65.6	7.3	65899	50	AC006436	AC006436 Arabidops
35	65.2	7.2	94362	8	AC005405	AC005405 BAC F18A1
36	64.8	7.2	127053	8	T1N24	AF149413 Arabidops
37	64.8	7.2	75568	50	AC006532	AC006532 Arabidops
38	64.6	7.2	97711	8	ATT16K5	AL132965 Arabidops
39	64	7.2	104204	8	ATT9C5	AL132964 Arabidops
40	64	7.1	114935	5	AC007651	AC007651 Arabidops
41	63.4	7.0	3541	8	A67432	A67432 Sequence 5
42	63	7.0	3541	8	AF053997	AF053997 Lycopersi
43	63	7.0	6695	5	A67796	A67796 Sequence 1
44	63	7.0	3979	5	A67428	A67428 Sequence 1
45	62.8	7.0	3979	5	A67428	A67428 Sequence 1

ALIGNMENTS

```

RESULT 1
LOCUS A67823 894 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 28 from Patent WO9743427.
ACCESSION A67823
VERSION A67823.1 GI:4756645
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 894)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICETIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source location/Qualifiers
1..894
/organism="unidentified"
/db_xref="taxon:32644"
<1..678
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42250.1"
/db_xref="GI:4756646"
/translation="GPIDQSEGDALHAIKRLSDPDNYVQSDPTLVNPTWHTYCN
OHQVTRLDLGNLSGHLVPELGRLEHLYLGLITLLPDDYIKRTTLLSVTHITCE
ESYSELVKNIEIGITIPSELGNLSLSIDLYNNMTGKPSLKLIVFLRNENR
LNGPIPRELTVISSLSKVDVSGNDLCGTPVEGPEPIHPQNFENNRLRGPPELLGIA
SYDNCT"
BASE COUNT 270 a 163 c 176 g 285 t
ORIGIN
Query Match 100.0%; Score 894; DB 5; Length 894;
Best Local Similarity 100.0%; Pred. NO. 7.9e-195;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ggacgattcaagctccgaaggagagcgtctcaagcgcttcgaggagctatcagat 60
Db 1 GGACCGATTCAAGCCTCGAAGGGAGCGCTTCACGCCCTCGCGGAGCTTATCAGAT 60
QY 61 ccagacaatgtgttcagagttgagatccaactctgttaactctgtacttggttcat 120
Db 61 CCAAGACATGTTGTTGAGAGTGGATGCCAACCTGTGTAATGCTGTACTGGTTTCA 120
QY 121 gtcaactgttaatacaaccatcaagtcactgcgtcgtgattggggaattcaactatct 180
Db 121 GTCACTGTGTAATCAACCAATCAAGTCACTGCTGTGATTGGGGAATTCAAACTTATCT 180
QY 181 ggacatctagaactgaactgggaagcttgaaacattacaatactgtatggaatcattc 240
Db 181 GGACATCTAGTACTGAAGCTTGGAACTTGAACATTATACAAATGCTGTAATGGAATCATC 240
QY 241 actcttgccttctgattatctgaaacaattacattatcagtcacacataaacattt 300
Db 241 ACTCTTGGCTTTGATTATCTGAATAACATTATACATTATACATCAATATTAACATT 300
QY 301 tgccttgatcatatagtgaaacttacaataaagaaagaaacttaacttctgag 360
Db 301 TGCCTTGATCATATAGTGAAGCTTACAAAACAGATTAAGAAAGAACTATACCTTTGAG 360
QY 361 ctggaatctgaaagagctaatcagttctgattctgacaacaacaactcaccagggaaa 420
Db 361 CTGGAATCTGAAGAGCTAATCAGTTTGATGATGTAACAACAATCTGCACCGGAGAA 420
QY 421 atccatcttctctggggaanaattgaagtcactgtttcttgcggcttaagaaacgga 480
Db 421 ATCCATCTTCTCTGGGGAANAATTGAAGTCACTGTGTTTTTGGCGCTTAACGAAACCA 480
QY 481 ttgacggctcattatccctagagaactcacagttattcagaagccttaagttgttgatgc 540
Db 481 TTGACGGCTCTATTCTCTAGAGAACTCACAGTTATTTCAGCCTTAAGTTGTGATGTC 540

```

```

QY 541 tcaaggaaatgattgtgtggaacaattccagtagaaggaccttctgaacatctccatg 600
Db 541 TCAGGAATGATTGTGTGGAAACAATTCAGTAGAAGGACCTTTTGAAACATATTCCTATG 600
QY 601 caaaacttgggaacaactgagattggaggagccgaactataagttcttcgagactat 660
Db 601 CAAAACCTTGGAAACACCTGAGATGGAGGACCAAGAACTACTAGGCTCTGGAGACTAT 660
QY 661 gacacaaattgacttaaaaagaagttgaagaacctataaagaagaattgtagtacct 720
Db 661 GACACAAATTGACTTAATAAAGAGTTGAAGAACCTATAAAGAAGATGTAGTGACCT 720
QY 721 tgaagaactctgtacaaagtgtttgtaaatcatatagaagcctgtttcattatata 780
Db 721 TGAAGAACCTCTGTACCAAGTGTGTGAATCAATATAGAGCCCTTGTTTCATGTTATATA 780
QY 781 tgaagacttggagagacagtaactgtcaatgtatgtgttgtagaanaagttgaatg 840
Db 781 TGAAGCTTTGAGAGACAGTAACTTGCAATGATTTGGTATTTGGTAAAGAAAGTTGAATG 840
QY 841 agaattgcttgaattgaattgattgttcttattgtaactgaaattcttata 894
Db 841 AGAATTGCTTTGTAATGGAATTTGTTTCTTATGTAATTTGAATTTCTATATA 894

RESULT 2
LOCUS A67821 789 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 26 from Patent WO9743427.
ACCESSION A67821
VERSION A67821.1 GI:4756643
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 789)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICETIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source location/Qualifiers
1..789
/organism="unidentified"
/db_xref="taxon:32644"
<2..664
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42248.1"
/db_xref="GI:4756644"
/translation="RPTRPARYRWELFPAASLIITLAIHLIVANSSEGDALVALRSLT
DPDHYLGSDPTLVNPTWPFVTCNONRYTRIVDLNSNSGHLAELEKLEHLYLE
LYKNIOGTTIPSELGNKNLSIDLYNNMTGIVPSLGLKSLVFLRNDRLTGPI
PRALTAIRSLKVDVSSNDLCGTPINGPRAHILPIQNFENRPLRBEPELLGLASVDIN
CT"
BASE COUNT 235 a 191 c 156 g 207 t
ORIGIN
Query Match 40.4%; Score 361; DB 5; Length 789;
Best Local Similarity 72.8%; Pred. NO. 6.9e-73;
Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2;
QY 15 ctccgaaggagagcgtcttcacagcgcttcgagagctatacagatccaagaatgtgt 74
Db 91 CTCGGAAGAGAGTCTCTTTAGCGCTCTTCGCCGAGTTTAACAGATCCGAGCATGTCT 150
QY 75 tcaagattggatccaactcctgttaactcctgtacttggtttcatgtcactgtataca 134
Db 151 CCAAGCTGGGATCCAACCTTGTTAATCTGTACCTGGGTTCATGTCACCTCTAACA 210
QY 135 aacccatcaagtaactcgtctgatttgggaattcaactatcctgagacatcagtagacc 194

```

Db 211 AGACAAACCGGTCACCTGCTGATTTGGGAATTCAAACCTCTGTGACATCTTGCCGC 270

QY 195 tgaacttgggaagcttgaacatttacaatattgtatgaatcatcatttgcctt 254

Db 271 TGAGCTTGGGAACCTGGAACATTTCAGATATC----- 304

QY 255 tgattatcgtaaaaacatttacaattatcagtcacacataataacatttgccttgatcata 314

Db 304 ----- 304

QY 315 tagtgaactctacaaaacagagattcaaggaactatccttgaagcttgaacattgaa 374

Db 304 ---AGAGCTTACAAAACAAACATCCAAAGGAACTATACCTTCCGAACCTGGAAATCTGAA 360

QY 375 gagtcaatcagtttggatctgtacaaacaacatctcacgggaaatccacttctt 434

Db 361 GAATCTCATACGCTTGATCTGTACAAACAACATCTTACAGGAAATGATCCCACTCTCTT 420

QY 435 gggaaaattgaagctactgttttttttggggttaacgaaacggaatgacgggtccat 494

Db 421 GGGAAATTTAAGTCTCTGTCTTTTACGGCTTAATGAAACCGATTACGGGGCCAT 480

QY 495 tcttagagaactcacagattatccaagccttaagttgttgaatgtctcagggaaatgatt 554

Db 481 CCTTAGAGCACTCATCTGCATCCCAAGCCCTTAAGTTGTTGATGCTCAAGCAATGATTT 540

QY 555 gtgtggaacatccagtagaaggaacctttgacacatcctctatgacaaacttggaa 614

Db 541 GTGTGGAACAATCCCAACAACGACACTTTGTCTCATCTCTTACAACTTTGAGAA 600

QY 615 caacttgaagttggaggacacgaactaagttcttggagctatgaacacaaattgac 674

Db 601 CAACCCGAGGTGGAGGACCGGAATTAATCTCGGTTCGCAAGTACGACATCTAATGAC 660

QY 675 ttaaaaagaagttggaagaacctataaagaaga-tgttgggtgacctgttgaagacctg 733

Db 661 CTGAAAAAATTTGGCAAAACCTGAAATGAGAAATGGGGGTGACCTGTGAAACACTT 720

QY 734 taccagtgcttgaatc 752

Db 721 CACCACCTTTATCAAAATATC 739

RESULT 3

A67825 1063 bp DNA PAT 05-MAY-1999

LOCUS A67825

DEFINITION Sequence 30 from Patent WO9743427.

ACCESSION A67825

VERSION A67825.1 GI:4756647

KEYWORDS

SOURCE

ORGANISM

unidentified.

unclassified.

REFERENCE

1 (bases 1 to 1063)

AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.

TITLE PRODUCTION OF APOMICRITIC SEED

JOURNAL Patent: WO 9743427-A 20-NOV-1997;

CTBA GEICY AG (CH)

FEATURES

Location/Qualifiers

1..1063

/organism="unidentified"

/db_xref="taxon:32644"

106..762

/note="unnamed protein product"

/codon_start=1

/protein_id="CAB42252.1"

/db_xref="GI:4756648"

/translation="MASRYRWELFASLITLALIHVEANSEGDALYALRSITDP

DHYLGSDPLVPCWTFVTCNDNRVTVGDSNLSGHLAPGLGHEHQYLEIK

KNNIOGTIPSELGNLKLNSLIDLYNNNLGIYVTSGLKSLVFLRLDNRITGPIR

ALTAISLKVVDVSSNDLGLTPTNPFHPIPLQNFENNPRLEGEELGLASDTNCT

BASE COUNT 313 a 242 c 206 g 302 t

ORIGIN

Query Match 40.4%; Score 361; DB 5; Length 1063;

Best Local Similarity 72.8%; Pred. No. 6,76-73;

Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2;

QY 15 ctccgaagggaagcgtctcagcgcttcgcggagctatcagatccagaatgtgt 74

Db 189 CTCGGAAGAGATGCTCTTACGCTCTTCGCGGAGTTAAACATCCGACATGTTCT 248

QY 75 tcsagttggatccaaccttggtaatcccttgaactgtgttcaatgaacttaata 134

Db 249 CCAAGCTGGAGATCCACCTTGTAAATCTGTAACTGGTTCATGTCACCTGTAAACA 308

QY 135 acacccatcagtcactcgtctgtgatttgggaattcaaatcatctggaactagacc 194

Db 309 AGACAACCGGTCACCTCGTGTGATTTGGGAAATTCAAACCTCTGTGACATCTTGCGCC 368

QY 195 tgaacttgggaagcttgaacattacaatctgtatgtaataatcatcacttgcctt 254

Db 369 TGAAGTTGGAGACTTGAACATTTACAGTATCT----- 402

QY 255 tgattatcgtaaaaacatttacaattatcagtcacacataataacatttgccttgatcata 314

Db 402 ----- 402

QY 315 tagtgaactctacaaaacagagattcaaggaactatcactctcgtgacttgaacattgaa 374

Db 402 ---AGAGCTTACAAAACAAACATCCAAAGGAACTATACCTTCCGAACCTTGAAATCTGAA 458

QY 375 gagtcaatcagtttggatctgtacaacaacatctcccggaatcccatcttctt 434

Db 459 GAATCTCATACGCTTGATCTGTACAAACAATCTTACAGGAAATGATCCCATCTCTT 518

QY 435 gggaaaattgaagctactgttttttggcgcttaacgaaacggaatgacgggtccat 494

Db 519 GGGAAAAATTTGAAATCTCTGTCTTTTACGGCTTAATGAAACCATTTGAGGGGCCAT 578

QY 495 tcttagagaactcacagttatcttcaagccttaagttgttgaatgtcgaaggaaatgtt 554

Db 579 CCTTAGAGCACTCATCTGCATCCCAAGCCTTAAGTTGTTGATGCTCAACAAATGATTT 638

QY 555 gtgtggaacaaatccagtcagtcgaaggaaccttgaacacatctctatgcaaaacttggaa 614

Db 639 GTGTGGAACAATCCCAACAACGGAACCTTTGCTGCATCTCTTACGAACTTTGAGAA 698

QY 615 caactgagatttggagagccgaactaagttcttgcggactatgacacacattggac 674

Db 699 CAACCCGAGGTGGAGGACCGGAATTAATCTCGTCTTGCAAGCTACGACTTAACCTGCAC 758

QY 675 ttaaaaagaagttggaagaacctataaagaaga-tgttgaagttgacctgttgaagacctg 733

Db 759 CTGAAAAAATTTGGCAAAACCTGAAATGAGAAATGGGGGTGACCTGTGAAACACTT 818

QY 734 taccagtgcttgaatc 752

Db 819 CACCACCTTTATCAAAATATC 837

RESULT 4

A67819 981 bp DNA PAT 05-MAY-1999

LOCUS A67819

DEFINITION Sequence 24 from Patent WO9743427.

ACCESSION A67819

VERSION A67819.1 GI:4756641

KEYWORDS

SOURCE

unidentified.

unclassified.

REFERENCE

1 (bases 1 to 981)

AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.

TITLE PRODUCTION OF APOMICRITIC SEED

JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES Location/Qualifiers
source 1..981
/organism="unidentified"
/db_xref="taxon:32644"
CDS 104..760
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42246.1"
/db_xref="GI:4756642"
/translation="MASRMYRELPAASLTTLALIHVEANSSEGDALVALRSLTDP
DHYLOSMDPTLVNPFCTWFHYTCNODNRVTRVDGNSNLSGHLAPBLGKLEHQLYELY
KNNIQGTIPSELGNLNLISLDIYNNNLGIYPTLGIKLSLVFRLNDNRITGPILR
ALTAIPSLRVYDVSSNDLGGTITPTNPFPHIPLQNFENNPRLGPELLGLASYDINCT
"

BASE COUNT 286 a 236 c 180 g 279 t
ORIGIN

Query Match 39.8%; Score 356.2; DB 5; Length 981;
Best Local Similarity 72.4%; Pred. No. 8.5e-72;
Matches 535; Conservative 0; Mismatches 113; Indels 91; Gaps 2;

QY 15 ctccgaaggagagcgtcttcacgcgcttcgcgaggtatcatgaccagacaattgtt 74
DB 187 CTCGGAAGGAGATGCTCTACGCTCTTCGCCGAGTTGACAGATCCAGACATGTCCT 246
QY 75 tcaagagttggatccaactcttgtaactccttgacttggttcaatgctgtaacca 134
DB 247 CCAGAGCTGGGATCCAACTCTTGTTATCCTTGACCTGGTCCATGCTACCTGTACCA 306
QY 135 acaccatcaagtcactcgtctgctggtttggggaattcaaatctcggacaattagacc 194
DB 307 AGACAACCGCGCTCACTGCTGTGGATTGGGAAATCTTAACCTCTGACATCTTGCGCC 366
QY 195 tgaacttggagagcttgaacattatcatatctgtatgaaatcatcactcttgcctt 254
DB 367 TGAGCTTGGAAGCTTGAACTTACAGTATCT----- 400
QY 255 tgattatctgaaaaacattatcatatcatgacacacataacatttgccttgatcata 314
DB 400 ----- 400
QY 315 tagtgaactctacaacaaacgagattcaaggaactatacctcttgagcttggaaatcgaa 374
DB 400 ---AGAGCTTACAAAACAAACATCCAGAGACTATACCTTCGAACTTGGAAATCTGAA 456
QY 375 gagcttaactcagtttgatctgtacaacaaatctcacccgggaaaaatcccatcttctt 434
DB 457 GAATCTCATCAGCTTGATCTGTACAAACAAATCTTACAGGAGTATGTTCCCATCTCTT 516
QY 435 gggaaaattggaagtcactgttttttgcggttaacgaaacccgattgacgcgtctat 494
DB 517 GGGAAAATTGAAGTCTCTGCTTTTACGGCTTAATGACACACGATTTGACCGGCTCAAT 576
QY 495 tccctgaactcacaagatatttcaagccttaagttgttgatcttcaaggaattgatt 554
DB 577 CCTTGAAGCAGCTCAAGGGAATCCCAAGGCTTAAGTTGTTGAGCTCTCAAGCATGATT 636
QY 555 gttgtggaacaattccagtagaagaacctttgaaacacattccctatgcaaaacttgaa 614
DB 637 GTGTGGAACAATCCCAACAAACGAGCCTTGTGCTACATCTCTTACAGAACTTGGAA 696
QY 615 caacttgatattggaggagacagaactactagtccttgcgactatgaacccaattgac 674
DB 697 CAACCCGAGATTGGAGGACCGGAATTACTGCTTTGCAAGCTACGACATTACTCAC 756
QY 675 ttaaaaagaattgagaagacctataagaaga--tgtaagttgacacttgtaagaagctct 733
DB 757 CTGAACAACAATGGCAAAACCTGAAGAATGAGAAATGGGGGCTGACCTGTGAAGAACCTT 816
QY 734 taaccaagtgcttgtaataac 752

DB 817 CACCACCTTATCAAAATATC 835
||||| | | | |
RESULT 5
A67817
LOCUS A67817 1106 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 22 from Patent WO9743427.
ACCESSION A67817
VERSION A67817.1 GI:4756639
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1106)
AUTHORS De/V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES Location/Qualifiers
source 1..1106
/organism="unidentified"
/db_xref="taxon:32644"
CDS 142..798
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42244.1"
/db_xref="GI:4756640"
/translation="MASRMYRELPAASLTTLALIHVEANSSEGDALVALRSLTDP
DHYLOSMDPTLVNPFCTWFHYTCNODNRVTRVDGNSNLSGHLAPBLGKLEHQLYELY
KNNIQGTIPSELGNLNLISLDIYNNNLGIYPTLGIKLSLVFRLNDNRITGPILR
HSRQSAKFYVDVSSNDLGGTITPTNPFPHIPLQNFENNPRLGPELLGLASYDINCT
"

BASE COUNT 331 a 258 c 206 g 311 t
ORIGIN

Query Match 36.9%; Score 330; DB 5; Length 1106;
Best Local Similarity 71.9%; Pred. No. 8.3e-66;
Matches 532; Conservative 0; Mismatches 115; Indels 93; Gaps 4;

QY 15 ctccgaaggagagcgtcttcacgcgcttcgcgaggtatcatgaccagacaatggtt 74
DB 225 CTCGGAAGGAGATGCTCTACGCTCTTCGCCGAGTTGACAGATCCAGACATGTCCT 284
QY 75 tcaagagttggatccaactcttgtaactccttgacttggttcaatgctgtaacca 134
DB 285 CCAGAGCTGGGATCCAACTCTTGTTATCCTTGACCTGGTCCATGCTACCTGTACCA 344
QY 135 acaccatcaagtcactcgtctgcatcttggaattcgaactatctgacatctagtaacc 194
DB 345 AGACAACCGCGCTCACTGCTGTGGATTGGGAAATCTTAACCTCTGACATCTTGCGCC 404
QY 195 tgaacttgggaagcttgaacattatcatatctgtatgaaatcatcactcttgcctt 254
DB 405 TGAGCTTGGAAGCTTGAACTTACAGTATCT----- 438
QY 255 tgattatctgaaaaacattatcatatcatgacacacataacatttgccttgatcata 314
DB 438 ----- 438
QY 315 tagtgaactctacaacaaacgagattcaaggaactatacctctgagcttggaaatcgaa 374
DB 438 ---AGAGCTTACAAAACAAACATCCAGAGACTATACCTTCGAACTTGGAAATCTGAA 494
QY 375 gagcttaactcagtttgatctgtacaacaaatctcacccgggaaaatcccatcttctt 434
DB 495 GAATCTCATCAGCTTGATCTGTACAAACAAATCTTACAGGAGTATGTTCCCATCTTCT 554
QY 435 gggaaaattggaagtcactgttttttgcggttaacgaaacccgattgacgcgtctat 494
DB 555 GGGAAAATTGAAGTCTCTGCTTTTACGGCTTAATGACACACGATTTGACCGGCTC-AA 613


```
QY 495 tccatgaagactcacagattatccaagcc-ttaaagttgttgatgtctcagggaatgalt 553
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 TCCTAGAGACACTACGAGCAATCCCAAGCCTTTAAAGTTGTGACGTCTCAAGCAATGATT 673
QY 554 tgttgaagaatcccgtaggaagcgctttgaacacattccatgaacaaacttgaga 613
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 TGTGTGGACATCCCAACAAAGGACCTTGTCTCATTCTTACAGAACTTGAAGA 733
QY 614 acaactgagattgagagaccagaactactagctctcgagactatgacacaaatgca 673
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 ACAACCCGAGATTGAGAGACCGGAATTACTCGCTGTGCAAGCTACGAACTGCA 793
QY 674 cttaaaagaagttgagaacctataagaaga-tgttaggtgacctgtgaacctt 732
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 CCTAAACAATGGCAAAACCTGAAATGAAGATTGGGGGTGACCTTGAAGAACT 853
QY 733 gtaccaagttgttaaatc 752
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 TCACCACCTTATCAAAATATC 873

RESULT 6
LELRGENE 4604 bp DNA PLN 10-OCT-1996
LOCUS L.esculentum LRP gene.
DEFINITION X95269
ACCESSION X95269.1 GI:1619299
VERSION LRP gene; LRP protein.
KEYWORDS LRP gene; LRP protein.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
Asteridiales; Solanales; Solanales; Solanales; Solanales; Solanales;
Lycopersicon.
REFERENCE 1 (bases 1 to 4604)
AUTHORS Vera, P.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1996) P. Vera, Universidad Politecnica de
Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino
de Vera 14, E-46022 Valencia, SPAIN
2 (bases 1 to 4604)
AUTHORS Tornero, P., Mayda, E., Gomez, M.D., Canas, L., Conejero, V. and Vera, P.
TITLE Characterization of LRP, a leucine-rich repeat (LRP) protein from
tomato plants that is processed during pathogenesis
JOURNAL Plant J. 10 (2), 315-330 (1996)
MEDLINE 96367673
FEATURES
SOURCE location/Qualifiers
1..4604
/organism="Lycopersicon esculentum"
/cultivar="VFN8"
/db_xref="taxon:4081"
/tissue_type="leaf"
<1..345
/number=1
113..4355
/gene="lrp"
join(113..345,2381..2452,2565..2708,2928..2999,3081..3152,
4283..4355)
/gene="lrp"
/codon_start=1
/gene="lrp"
/product="LRP protein"
/protein_id="CA64565.1"
/db_xref="GI:1619300"
/db_xref="SPTREMBL:096477"
/translacion="MEAVKTSLSFLKMGLLAVAVAVKNSGDALVALRSL
SDPGNVLSQSDPNLVNPGCTWFHTVTCNGDNGVREVDGNSKLSGHLVPELGLKLEHLYL
ELYNNINQGTIPKELNKSLSISLDYNNISGTSISGLNKNLVFLINDNKTGTP
IPRELTSISLKVVDVSNNDLCTIPTSGFEHIIPLANFHNHPRLEGPRLGIATYDI
NCS"
346..2380
/gene="lrp"
/number=1
2381..2452
exon
Intron
exon
```

```
Query Match 13.4%; Score 119.4; DB 7; Length 4604;
Best Local Similarity 65.7%; Pred. No. 1.1e-17;
Matches 216; Conservative 0; Mismatches 91; Indels 22; Gaps 2;

QY 156 ggatttgggaattcaacttactctgacatctagctgactggaacttgggaacttgaaca 215
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2381 GGATCTTGGGAACCAAGTATGCTGTCATTTGGTACCTGAGCTCGGAACCTTGAACA 2440
QY 216 ttacaatctgtagtgaatc-----cactcttgccttgcattattctg 264
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2441 TCTACATATCTGAGATATTTTGAATCTGAAATGACTTGACATTTGACTTTT 2500
QY 265 aaacattcatatc-----agtcacatatcaacttgccttgatcat 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2501 TAGCAATTAATCTTATTGTTTAAAGAGTCAATCAATGAGTAAATGTTTGTCT 2560
QY 314 atagtgaacttcaaaaacgagattcaaggaactatcctctcgtgagcttgaacttga 373
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2561 TCAGGAGCTTTACAAAATAATATTCAGGAGCATCCCTTAAGAGCTCGTAACCTTGA 2620
QY 374 agagctaaatcagtttgatctgtacaacacatctcaccgggaatcccatcttct 433
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2621 AGAGCTTATAGCTGATGATCTGACACACATATTTTCGGGCAATTCCTACTTCAAC 2680
QY 434 tgggaatgaagtcacgtgtttttt 462
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2681 TTGGAACCTGAAAAAACCCTTGTTCCTTG 2709

RESULT 7
LELRGENE 2089 bp DNA PAT 05-MAY-1999
LOCUS A67827
DEFINITION Sequence 32 from Patent WO9743427.
ACCESSION A67827
VERSION A67827.1 GI:4756649
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE 1 (bases 1 to 2089)
AUTHORS De, V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
```


DEFINITION Sequence 2 from Patent WO9743427.
ACCESSION A67797
VERSION A67797.1 GI:4756623
KEYWORDS
SOURCE carrot.
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
REFERENCE
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMITIC SEED
JOURNAL Patent: WO 97/43427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source location/Qualifiers
1..1815
/organism="Daucus carota"
/db_xref="taxon:4039"
94..1755
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42241.1"
/db_xref="GI:4756624"
/translation="MNRSINILNYMFTDAYLDKYGVLMTLEYSNNISGPIPSDLG
NLTVSLDILVNSPSPGIPDLGLRLRLRLNNLSGPIPSMPTLITLOVLD
SNRLSGVPDNGSFLPTISPANNILGPTVGRPCSPSPSPPTPIPTOP
PGONGPTGAIAGVAGAGALFPAPAPAFMRPRRPFVPAEDPEVLOGLK
RRLURELVADITSTIIGRGFGVYKGRADSLVAVKRLKERTPEGELOFTEV
EMISAVHNLRLNGFCMPTERLIVPYANGSVASCLREKOPSPPLMDPTRRRI
ALGSAARGSLYHDHCDPKIHRDVKANILDEFEAVVDGFLARLMDYKDHVTA
VAGTGLYLAPEYLSGKSEKTDVFGVIMLELITGQRAFDLARLMDVMDLVV
KSLKEKLEMLVDPLENNYIDTEVEEDLVALCTGSPMERPKSEVVRMLEDIDG
LAERKDEMOKEVIVHDOVELAPHRPRTSEWILDSTONLAFLSLGPR"
BASE COUNT 531 a 354 c 415 g 515 t
ORIGIN
Query Match 10.7%; Score 95.4; DB 5; Length 1815;
Best Local Similarity 56.1%; Pred. No. 3.7e-12;
Matches 180; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
QY 319 gaactctacaagaaggttcaaggaactatccttgcgcttggaacttggaaggt 378
DB 178 GACCTTTACGCAATTAACCTAAGTGACCAATTCCTAGGATCTTGGGAATCTGCAAAAT 237
QY 379 ctaacagtttgatctgtacacaacaatctcacccgggaaatccatctcttgga 438
DB 238 TTGGTAGCTTGACCTATACATGATAGCTTCTGTGACCTATACCGGACACATTAAGA 297
QY 439 aaattgaagtcactgttttttttggcgcttaacgaaacggaattgaacggtccattctc 498
DB 298 AACCTTACAAGGCTAAGATCTTGGCTTCACACACAAACAGCCTCTGCGTCCAAATTC 357
QY 499 agagaactcacagttattcaagccttaagttgttgatgtctcaggggaatgattggt 558
DB 358 ATGCACTGACTAATATTAACAACCTTCAAGTCCGGAATTTACAACAATGCGGTATACA 417
QY 559 ggaacaattccagtagaaggaacttttgacaacattcctatgcaaacatttgagaacaac 618
DB 418 GGAACGCTACCGGATTAATGAGCTATTTCTTTTACACCTATCAGTTTGCACAAATAT 477
QY 619 ctgagattggaaggaccagaa 639
DB 478 TTGAATTATATGTGACCCGTA 498
RESULT 10
SB062279 936 bp mRNA PLN 09-DEC-1996
LOCUS
DEFINITION Sorghum bicolor leucine-rich repeat-containing extracellular
glycoprotein mRNA, complete cds.
ACCESSION U62279
VERSION U62279.1 GI:1710123

KEYWORDS
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum.
REFERENCE
AUTHORS Hipskind, J.D., Nicholson, R.L. and Goldsbrough, P.B.
TITLE Isolation of a cDNA encoding a novel leucine-rich repeat motif from
Sorghum bicolor inoculated with fungi
JOURNAL Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)
MEDLINE 97124217
ERRATUM: [[published erratum appears in Mol Plant Microbe Interact
1997 Mar;10(2):302]]
REFERENCE
AUTHORS Hipskind, J.D.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant
Pathology, Purdue University, West Lafayette, IN 47907, USA
FEATURES
source location/Qualifiers
1..936
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone="SLRR"
12..755
/function="protein-protein ligand binding"
/note="leucine-rich repeat-containing extracellular
glycoprotein; contains six N-glycosylation sites
[NX(S/T)]"
/codon_start=1
/protein_id="AAC49559.1"
/db_xref="GI:1710124"
/translation="MAPQFAAGFLNGILATFAFASCTESGDIYKORLADPEPNVL
OSWSTLANCTPEPHYVQNNNNVIVIVDVGNGNISRLPLDAEQNLQVITLNG
NGSIPETLGNLNLISLDLMDNLITGEIPTTSSSTLRLKIQNNLAGPIPSFGN
LTSLESKLQENSGAIPASLGNKALPRLNDMLNGTVPSKSFPLTFGNLTEL
NTDRNNDGRTSSGLNVAIIODALKTA"
BASE COUNT 247 a 232 c 207 g 250 t
ORIGIN
sig_peptide
Query Match 10.4%; Score 93; DB 8; Length 936;
Best Local Similarity 51.8%; Pred. No. 1.4e-11;
Matches 250; Conservative 0; Mismatches 215; Indels 18; Gaps 1;
QY 15 ctccgaaggggcgctcttcacggcgtccgcggaagcttatcgatccagaatgtgt 74
DB 83 CACTGAGGTGACATACGTACAGCAAGGTTGCAATGGAGAGACCAACAACGTGCT 142
QY 75 teagaattggatccaaactctgttaactctgtactgttctcactgttaact 134
DB 143 GCAAGACTGGAATTCGACGCTGCCAATCCTGCACCTGGTTCATGTCACTGCAACAA 202
QY 135 aacaccaatgaatcagctgtctggaatttggaatcaacttaactcgtgaacattagc 194
DB 203 CAATTAATCTTGTATCCCGGTGGATTGGAATGCAAGGCAATCTCGGTCTGTGCTTC 262
QY 195 tgaacttgggaagcttgacaattcaatatactgtatggaatcaatcactctttgcttt 254
DB 263 AGATCTTGCAAAATTCGAACCTCCAGTACATGAGGTGATGCAAGCGCTTGAAACG 322
QY 255 tgaattatcgaataaatttaactatctacgtcaacataataacatttgctttgaatcata 314
DB 323 TTGCATTCGAGAAACACGAGGCAACCTGACATCACTCATGAGCTT----- 368
QY 315 tagtgaactctacaagaagcttgcaaggaactatacctctcgaattggaatctgaa 374
DB 368 ---GGATCTCTGGGCAACACTTCTTACCGGGAATGCCAATCAGCTTGTTCTGTCTGAG 424
QY 375 ggaattcaatcagtttgatctgtacaacaacattcgaacgggaataatccatctcttt 434

Db	425	CAGCGCATATNTGAGTGTACCAAAACAACCTGCAGGGCCATTATCATTACATCGTT	484
OY	435	gggaanaattgaagtcactctgttttccgcttaagaagaacagtatgcgcgtctc	494
Db	485	TGGCAACTGACTACGCCCTTCCTTGGAATCGAAGCTTCAGAGAAATTCGTTAGCGCGCCATN	544
OY	495	tcc 497	
Db	545	TCC 547	
<hr/>			
RESULT	11		
LOCUS	TSF17/c		
DEFINITION	TSF17	43277 bp	DNA
ACCSSION	AL049917		
VERSION	AL049917.1	GI:4884021	
KEYWORDS	HTG; HTGS; PHASEL.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
JOURNAL	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
REFERENCE	Arabidopsis.		
AUTHORS	1 (bases 1 to 43277)		
TITLE	Lennard,N., Quail,M., Harris,B., Rajandream,M.A. and Barrell,B.G.		
JOURNAL	Direct Submission Submitted (20-MAY-1999) EU Arabidopsis sequencing project, The Sanger Centre, Wellcome Trust Genome Campis, Hinxton, Cambridge CB10 1SA, UK		
COMMENT	IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. For more information about this sequence or the ESSAIL project, see MIPS http://websvr.mips.biochem.mpg.de/proj/thal/. Order of segments is not known; 800 n's separate segments. NOTE: This is a 'working draft' sequence. This record will be updated with the finished sequence as soon as it is available and the accession number will * be preserved.		
FEATURES	Location/Qualifiers		
source	1..43277		
BASE COUNT	13207 a 7883 c 8609 g 13578 t		
ORIGIN			
Query Match	8.8%; Score 79; DB 32; Length 43277;		
Best Local Similarity	57.5%; Pred.No.1.6e-08;		
Matches 142; Conservative	0; Mismatches 105; Indels 0; Gaps 0;		
OY	329	aaacagagataccaaggatactacctctcgagcttggaacaatgaagagtcataatcagtt	388
Db	30596	ACAACGAAATTCAAAGATCGCATTCGATTCACACGAGATTGTGAACAATAACGAGCTCAAGTAC	30537
OY	389	tggatcggtcaaacacatctcacaccgggaagaatccatctcttttgggaatatgaagt	448
Db	30536	TTCATTTCGCCATCGGAAGCTCAAGTGGAGATTCCTCACAGACTTGGAGAAGCTCAAAGT	30477
OY	449	cacttgttttttcgagcttaacgaaaacagattacccgagtcctatctcctaagaaatca	508
Db	30476	CACTCGAGAGCGCTTCCTCGTAGAGAAATAAATTCACCCGGAACAATTCGAAGAGATTA	30417
OY	509	cagctatttcaagcccttaaagttgttgatgcttcgcgggaatgatcttctgtggaacaattc	568
Db	30416	GAAACATTACACAGCTTAAGAGTTCTGATTTCTCCGATTAAGCGTTAACCGCGGAGATTTC	30357

Ox	569	cagtaaa 575	11	11	11
Db	30356	CANTGGA 30350			
RESULT	12				
LOCUS	AB029327				
DEFINITION	AB029327 2786 bp mRNA				25-DEC-1999
ACCESSION	AB029327.1	GI:6635235			
VERSION	AB029327				
KEYWORDS	elicitor-inducible receptor-like protein EIR, complete cds.				
SOURCE	Nicotiana tabacum mRNA.				
ORGANISM	Nicotiana tabacum				
REFERENCE	Nicotiana, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.				
AUTHORS	Takemoto, D. and Kawakita, K.				
TITLE	Isolation of the gene for EIR, an elicitor-inducible receptor-like protein, from tobacco by Differential Display				
JOURNAL	Unpublished (1999)				
REFERENCE	2 (bases 1 to 2786)				
AUTHORS	Takemoto, D. and Kawakita, K.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Daigo Takemoto, Nagoya University, Plant Pathology Laboratory, School of Bioagricultural Sciences, Chikusa, Nagoya, Aichi 464-8601, Japan (E-mail: d192006d@nobi.nagoya-u.ac.jp, Tel: 81-52-789-4031 (ex. 4031), Fax: 81-52-789-5525)				
FEATURES					
source	Location/Qualifiers				
gene	1..2786				
CDs	/organism="Nicotiana tabacum"				
	/db_xref="taxon:4097"				
	1..2586				
	/gene="e1r"				
	1..2586				
	/gene="e1r"				
	/codon_start=1				
	/product="elicitor-inducible receptor-like protein EIR"				
	/protein_id="BA88636.1"				
	/db_xref="GI:6635235"				
	/translation="MMWVKGSLFHLFFLFLFCLFVTFSTRTKATALLKKATLONS				
	NLSLYSPSSKACKRWGYCVCNGKVSGLDIPAGVIGITLNNPFSPLPVLIDS				
	KNQUGSLPEPLGKLTINTLVDSFNQSLDIPPGQSLAKQTHIDHLNLSISG				
	ELGHLRSVTELDISINTLVGSLPSSGSLNHLNLSLQCLKNNISGTHPEEIGVSLIQ				
	IDLNTLVNGSLIPASLENNLNLNLVLYENQSGSLPEIQLRTLDIRLNTNVLG				
	SLSPALGSLSLTIQLEHNLQSGSLPEEIGVRLVAVSLVTNLSGSLPSIGNTV				
	SLSGLSPENLISGPISSGLNDNLVLYLYLANQSGPISEGLNKLNLKMLADN				
	QNGSLIPFENLNNQVLFLENNNLGEPISLNNLSLKLVLGSRSLSGDTLQCL				
	INISLQVLTIPNNLSSELPSSICNLVSLRLIDLSRNNKLSIPQCGDMGSLVPL				
	DIHKNLSGTLPTPEPFGVSLRSLFHLHNNLSGKIPRLANKEILOYVDNLNLT				
	PMWGLTIPKLVQLRSLKSLISIRTSKDENNLELRINISVATLGNIPSLPQ				
	LKAMKIDIVLEKPYLGKFGADIRENLSVYTGKLEKIVRLVYITIDLSNR				
	FEGLHPSLNGELIALRVNLNRNGLOGHLPSSGLNLFVIESLDSFNQSGELPQQA				
	SQTLAVLNLSTYNNHLCIPQGFPHFENNYSYENGNLRFPPISCKCGDRVSEIN				
	NVTSVLQDSSEFLNDFKALMLVGSGLCIGLSILYFMISTGKLWLSRITEMQL				
	NFFII"				
BASE COUNT	818 a 520 c 502 g 946 t				
ORIGIN					
Query Match	8.7%; Score 77.8; DB 7; Length 2786;				
Best Local Similarity	57.0%; Pred. No. 3.8e-08;				
Matches 142; Conservative	0; Mismatches 107; Indels 0; Gaps 0;				
Ox	323	tctcaaaaacggagatcaaggaactactcttgagcgcttggaagtcaagagctaa 382			
Db	950	TGTAAGTAACTTCTTAATGAGTCTATATCCATATTCATATGGGGAATATGACAGCTTGT 1009			

OY	363	tcagtttggatctgtacaacaacaatctccacgggaaaaatccatctcttcttgggaaat	442
Db	1010	CCAGTTGTCCTTATATGAAATCATCTTCTTGCGCCCAATTCCTTCTGTTAGGGAAT	1069
OY	443	tgaatcactctgttttcttggcgcttaacggaacacgatgacggctctcttctcgaag	502
Db	1070	TGGACAACTTGCTGTAATTTGTAATCTTTATTCCTAACCAACTTCTGCTCCATCTCA	1129
OY	503	aactcacagattattccaagccttaaaagttgtatgctatgctcaggaatgattgttgaa	562
Db	1130	AATTAGGAATTTGAAAAACCTTAATATATGAAATGTCATGATCATCACTCAATG	1189
OY	563	caatccag	571
Db	1190	CGATTCCTG	1198
RESULT	13		
LOCUS	AP000815/c		
DEFINITION	AP000815	142418 bp	DNA
ACCESSION	AP000815		PLN
VERSION	AP000815.1		04-DEC-1999
KEYWORDS	GI:6498418		
SOURCE			
ORGANISM	Oryza sativa (cultivar: Nipponbare) DNA, clone: P0003H10.		
	Oryza sativa		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;		
	Poaceae; Oryza.		
REFERENCE	1 (bases 1 to 142418)		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC		
	clone:P0003H10		
JOURNAL	Published Only in Database (1999) In press		
REFERENCE	2 (bases 1 to 142418)		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-NOV-1999) to the DDBJ/EMBL/GenBank databases. Takui		
	Sasaki, National Institute of Agrobiological Resources, Rice Genome		
	Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan		
	(E-mail:ttasaki@agr.affrc.go.jp,		
	URL:http://www.dna.affrc.go.jp:82/, Tel:81-298-38-7441,		
	Fax:81-298-38-7468)		
COMMENT	The orientation of the sequence is from T7 to SP6 of the PAC clone.		
	Genes were predicted from the integrated results of the		
	following:GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as		
	SplicePredictor (October 1998 version). The genomic sequence was		
	searched against the non-redundant database NRP (PIR, SWISSPROT,		
	GENPEP, PDB) from MAFR DNA bank and the cDNA sequence database at		
	RGF. Protein similarities of the coding regions were searched		
	against NRP with BLASTP2.0. ESTs represent the identified cDNA		
	sequences using BLASTN 2.0 with the corresponding DDBJ accession		
	no. and RGP clone ID.		
	Detailed information on assemble quality together with annotation		
	of this entry at http://www.dna.affrc.go.jp:82/genomicdata/genomefil		
	nished.html.		
FEATURES			
source			
	Location/Qualifiers		
	1..142418		
	/organism="Oryza sativa"		
	/cultivar="Nipponbare"		
	/db_xref="taxon:4530"		
	/chromosome="1"		
	/clone="P0003H10"		
	join(12901..2951,3129..3317,3521..3850,4405..4638,		
	5069..5509)		
	/note="Similar to Arabidopsis thaliana DNA chromosome 4,		
	BAC clone F22K18 (AL035356)"		
	/codon_start=1		
	/protein_id="BA87822.1"		
	/db_xref="GI:6498419"		
	/translation="WGEPOKRLRYVWOGSNKFLCGRLIFGPAGSLFTSLTIVAAE		
	LVGLCCGCTTKMNSISSKQVYGLVTLATITVLGLADIAFLMTSSRDPGIVPRNAR		
	PLSCGCGEGEAGVDTPTPSAEWVAAPSHLTPSKYVNVNAGCVKRYCDTCLITY		
CDS			

CDs

PPAASHCSICNNCKAKEDHCHQPVGQCIGLRNTRFFFLITSTSLCYVYFVSWLMIVAAKDNDSDLSKSAAGPGLSVLIIVTYFVSVPVGGVILVHLYMTSNQVTPNPNFYRDKRKNPNRAISINIAVEFCAGIIPPSMNNFSPVAGPLVEFVSTNGQCLPFRNGADLTGVGEKDLDEMGKNGCIIPAILRGLDYDEMEKNDSYVHIKIDRGAADAPDPMARMHNEDESTPSTVSHVNSERL"

join(9165, 9635, 9735, .9799, 9992, .10319, 10422, .10520, 10743, .11006)

/note="ESTs D47452(S12946), C23573(S10086), C19173(E10057), AU081294(E10057), AU032455(S10086) correspond to a region of the predicted gene.; Similar to pNL134. (U37437)"

/codon_start=1

/protein_id="BAA87823.1"

/db_xref="GI:5698420"

/translation="MASSAMELSLNPAARGLSAKPRVYSRRIVYERVAASSAARPAKPTGPKRKRKLTIOETLTPRVTYTDDEMERFMEINQNLNOEFPDLIOEERFVIMVQNHVFNANPEFAADKMEGRLQIIVPEFESRCSTAEFGSLTYELGRKINIPVAALFMSLNSDEARHAGFLNKGISDNPALDGLTARAKTFEKFPEIFIAVYLS EKIGYRVLTIRHLKANPEQOYVPIKREYENNCODENRGEFSLAKAOEYFANV KALIMREFCLSVYVYMTLNDCCORTTPEEGIGLDEPMDHVIENRTTAIPPAVIVDNPERRKEDIRVYENIKKIITAIAGESDDIPVKNKRIPVHAALVSEITAIYAMPPTPESSVDPAEEPOLY"

join(13373, .13457, 15836, .15914, 17237, .17477, 17560, .17619, 17726, .17860, 18400, .18587, 18681, .18749, 18826, .18916, 19028, .19162, 19233, .19379, 19349, .19441, 19450, .19509, 19523, .19585, 19589, .19701, 19800, .19941, 20057, .20294, 21399, .21586)

/note="ESTs AU081256(C53656), AU081257(C53656), AU032236(R3784), AU081351(E61905), AU01387(E61905), AU081361(R3784) correspond to a region of the predicted gene.; Similar to alien-like protein. (AC005623)"

/codon_start=1

/protein_id="BAA87824.1"

/db_xref="GI:5698421"

/translation="MEDVGFESDDEBEDVDYENQYKSNVETPDEGLAEPDVRMEPKABMGFKALQVTKLYIKLGKIKMDAREMITIKSATVMSISEKINIMFVSSASQNSLSLOEFTYPTLKALEANEEMLAETKYLKATWEPMGYGRINSLIHLKHSIOECGSDQDKRGQLLEVALEIOWETENKNNKLAKLYLAKYALISIA PHRIKMIIECGGKMHAAEROMADAIDEFAEAKNDEAGNRPLOCYKYLNSM MESVNDPEQDEAKPYKNDPELTANTNLIAFYOKNDIMEEKLKSNRTINDPFIIT NYIEDLKIRNOVILIKIKPIYTRIRPIFSOKCKIISNMELIIVLFHGTLYVYRATIGSCSKYSKLDFLSCAOTRYQKLSCLNADPMFLPTEVNLISYVDALSSVLSAFCEMLPEKDEBOLVSLINDNITQGHIDNIEKDRDITPSHROVETSAE EHPENQGSWAHELAGCGCATGCTALGILVMMVQGCFLNLSNPSNVITLASSA ARNWEFMRHRSKFWYGRENWYEVKXVAPERTRTANIVLPSPSDVANOSISFONDEE KESFSLSTYDQOALNLOREHTTSLDNIINSESGYRPLAPSHSLGHPYTYTFVDDRDI"

CDs

join(22475, .22550, 23332, .23705)

/note="ESTs AU030013(E50493), AU081341(E50493) correspond to a region of the predicted gene.; Similar to O. sativa gene encoding calmodulin. (Z12898)"

/codon_start=1

/protein_id="BAA87825.1"

/db_xref="GI:6498422"

/translation="MADPDLOIAEFKASLPLDRKDGDCITIKELGVMSLQNNQ TEALQDMINEVDADGNGITIDPEFLNLMARKKQDSEBELKARVYDKQNGFITIAELIRHWNLSGKLTDEVEEMIREAVADSDGINDVEEVKAWAK"

complement(join(24778, .24897, 25615, .25722, 25808, .25893, 26090, .26177, 26435, .26638, 27519, .27581, 27708, .27778, 28339, .28389, 28512, .28552, 29058, .29225, 30476, .31136)

/note="hypothetical protein"

/codon_start=1

/protein_id="BAA87826.1"

/db_xref="GI:6498423"

/translation="MMFQSTSDARTHSGDGRSGAGRIACRGTRTSAAPEGLSREATIRGMALPEPASPRLVAGHRSRRRSPPSPBLVYTRACEEHRHRHRSCLPPIRGNARPPDPASPCCDPLATGVGRPMKAMDAAEPELVAAEENRAGAROLRGGGIVVNAHRSRGAEEERGGGFRSRRGVEEDKARVAGRENGKNGITIDBERMGSCHRSSTILMIQYTRARRGANNVLAHRCGDPRRRRSOCAGARRQALPIQSTFPLTYIMIDIKVAKAEDDITQFTAGYACTVAIASEVCRCHOALISITVSRRAALVAGVAGIGLSCETLHMHDKVEIDALEADADPISLGETTKSGGGRGHTSLKNNQMSATITVLCSESLHDVLTIFELFMSVSRKRGSLSPFSDGVGLVLTISGCVLYOALATVPLAKYVGIPIKPYFAVLSILLSTYPMMAVLYELEKVIININ ASLISNFMANKRMTNOSISLIDAGFHQRKSTLLSGSGYQKQOOGSKP"

CDs

CDS
complement(join(34468..34544,37202..37385))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA87827.1"
/db_xref="GI:6498424"
/translation="MAADKAGGDDAAAAAPLVSAAGRRRCPCGCTERCKADA
GIPYLNFEYIWWVCISFPFRPAKICDPSVYKMAHGHR"

CDS
join(40347..40358,40468..40680)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA87828.1"
/db_xref="GI:6498425"
/translation="MHPKRLHVAATVILVLOSPLPSSSFNNRRRRRQCPTAIAT
ADGSTRVLPDPDGGKGFRRMRKEGGD"
complement(join(42025..42083,43481..43661))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA87829.1"
/db_xref="GI:6498426"
/translation="MAGCKAGGDDAAAAAPLVPAAGRRRCPCGCTERCEAHG
IPYLNFEYIWWVCISAVNRYEATEPTSVS"
join(47081..47084,47194..47313,47469..47584)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA87830.1"
/db_xref="GI:6498427"
/translation="MSRGRAPATACGSRVCAEVAPEGDSARYALLLRDLP
AATPAKLRIRRDHGVPLREDDLSVCEPTD"
join(48412..48822,49072..49179,49283..49438,49876..50091,
50202..50411,50488..50625,50842..50943,51105..51170,
51274..51358,51458..51551)
/note="Similar to Arabidopsis thaliana chromosome II BAC
17A16 sequence; hypothetical protein. (AC005496)"
/codon_start=1
/protein_id="BAA87831.1"
/db_xref="GI:6498428"
/translation="MYLPLISVCSPSPAPSRRLDPDGGRGVDDLGVRVE
KRALAGLIVRSWTSSSRKGRHDDARDGTCDDGAGALLASVRYLLSGAQDD
AAEGEEDEGQGFPRMAIVFLCESAFILCMNDRNMSIALTPSAEAGNPQVGL
IOSGFVGLITLOAGIADIVGKTIVGFVWMSIATLTPPAKIGLPELIVTR
AFEGVEGAAAMPANNILSRKVPSESRSLALVSGMTVGLASPLINHFG
PVSFVSGSYGVWFSTMAKSAVSPLEDPGLSAEEKLITSGTGGEPVKIPGLI
LSKRPWALIVSHCHNMGTFILLTMPIYNOVKLNTEESGLCYVLPWMAVSAN
FGWADIVLSRGLSTYTRAKINOSIGFGRAPFLTQLSHDSPMAVYCAACSGCTD
AFQSQSLSHQDIPRIVAGVLGSLNTAGVAFGATATGTLIQHGSMDVFEVSY
VLYLVGTLVNLFSTEKIID"
complement(join(52616..52685,53228..53299,53395..53466,
53657..53800,54663..54734,56026..56234))
/note="ESTs C97644(C60871), A0081265(C60871) correspond to
a region of the predicted gene.; Similar to L.esculentum
LPR gene. (X95269)"
/codon_start=1
/protein_id="BAA87832.1"
/db_xref="GI:6498429"
/translation="MAAASPALAAVLAAVASANSSEGDALVALRRALADPGVLIQ
SMQPLTNPCTFHTACDRAGVTRLDIGASNSLGLAEPLGELHQLLEYKNNIO
GTLPAELSGSKNLISLDIVNNNTGTLPEIKGLSVLRINDNSLNPTRDLAKI
SGKALVDSNDKCGTIPPSGPEHPIPLNFDKNPLDEPQLGATLYTTC"
53608..53826
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA87833.1"
/db_xref="GI:6498430"
/translation="MKLIKDDHRSHTVFLVSLTSIALYTSLLTIYKGTGLAQPPE
KTGTSLLGVSYPESKDSRRKRPVYNLADH"
complement(join(63689..63842,63844..64145))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA87834.1"
/db_xref="GI:6498431"
/translation="MRSRARTDLRAAGRATATLRGRGRSRGRCQEVYTAGSGG
EGSDSWEBGRGGGGAARVAPAGERRRNRHAEKLTLLVGRGAPPHROSGGSH
EGGDEDEGELVGGALRPEEEEGGRGRGRGERRRRIVCREEEER

CDS

CDS

CDS

	Query Match	8.5%	Score 76	DB 7	Length 142418
	Best Local Similarity	65.1%	Pred. No. 7	3e-08	
	Matches 112	Conservative	0	Mismatches 60	Indels 0
				Gaps 0	
QY	298	ttttgctttgagtcataatgtaactctacacaaagagatcaagaaactaacctct	357		
Db	53620	TTTTTGAATAATTGTATACAGAGCTGTACAAAACATATTTCAGGGCATATCCAGCT	53761		
QY	358	gagcttggaaatcctgaagagctcaatcagtttgtaattctgtacacaaactccaccgg	417		
Db	53760	GACCTCGGGAGCTGTAGAAGACCTTTTGTGCTGGATCTGTATTAACAACAATTACTGGA	53701		
QY	418	aaatcccatctctcttgggaaatctgaagtcacattgttttttgcgacta	469		
Db	53700	ACCATACCGAAGAACTTGGCAGCTGAGTCTCTTAGATTTCTGTGTAGTAA	53649		

RESULT	14
AB010698	
LOCUS	AB010698 87434 bp DNA 20-NOV-1999
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MP1.12
ACCESSION	complete sequence.
VERSION	AB010698
KEYWORDS	AB010698.1 GI:2828183
SOURCE	HTG.
REFERENCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_11b:Mitsui Pl

ORGANISM
Arabidopsis thaliana
Eukarya: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; rosids II, Brassicales; Brassicaceae;
Arabidopsis.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (sites)
Nakamura, Y.
Structural Analysis of Arabidopsis thaliana Chromosome 5. V
Unpublished (1998)
2 (bases 1 to 87434)
Nakamura, Y.
Direct Submission
Submitted (28-JAN-1998) to the DDBJ/EMBL/GenBank databases.

Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2, 1532-3, Yana, Kisarazu, Chiba 292, Japan
(E-mail: ynakamu@kazusa.or.jp, Tel: +81-438-52-3935,
Fax: +81-438-52-3934)

FEATURES	Location/Qualifiers
source	1. .87434

```

/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MP112"
/clone.lib="Mitsui PI"
BASE COUNT      28656 a 15130 c 15592 g 28056 t
ORIGIN

```

Query Match	8.3%;	Score 73.8;	DB 7;	Length 87434;
Best Local Similarity	55.7%;	Pred. No. 2.4e-07;		
Matches 141; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0

323 tctacaaaacgagattcaaggaactataccttctgagcttggaatctgaaggtctaa 382

Db 52429 TGTATTCAATGATCTTGAAGGTCCAATTCCTGAAGAAATGTTGATATGAAGCTACTCT 52488

383 tcagtttgatctgtacaacaacatctcacgggaaaatcccatcttcttgggaaat 442

D5 52489 CAGTCTTGATCTTCCACACACAAATTCAGGTCAAATTCCTGCTCTGTCCTCAAGC 325348

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

QY 503 aactcacagttattcaagccttaagttgtagtctcaggaatgttgttgaa 562
Db 52609 GCCTTAGTCGCTTCGCTTCACACATTCGATATCCACATCTTCACTGAA 52668
QY 563 caattcacagtaga 575
Db 52669 CCATCCTGGAGA 52681

RESULT 15

E12705 3176 bp DNA PAT 24-JUN-1998
LOCUS Arabidopsis thaliana cDNA encoding a protein involved in
DEFINITION morphogenesis.
ACCESSION E12705
VERSION E12705.1 GI:3251537
KEYWORDS JP 1997056382-A/1.
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 3176)
AUTHORS Mitsuoka,N. and Kobatao,E.U.
TITLE GENE CODING FOR PROTEIN CONTROLLING MORPHOGENESIS OF PLANT
JOURNAL Patent: JP 1997056382-A 1 04-MAR-1997;
CHIKYU KANKYO SANGYO GIJUTSU KENKYU KIKO, MITSUI GYOSAI SHOKUBUTSU
BIO KENKYUSHO:KK

COMMENT OS Arabidopsis thaliana (thale cress)
PN JP 1997056382-A/1
PD 04-MAR-1997
PF 24-AUG-1995 JP 1995216187
PI MITSUKAWA NORIHIRO, ROBAATO EFU UITSUTEIA
PC C12N15/09,A01H5/00,C12N5/10;
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..3176
FT /organism='Arabidopsis thaliana' FT
FT CDS /strain='colombia',
FT 51..2981
FT /product='a protein involved in FT

FEATURES
source
morphogenesis'
Location/Qualifiers
1..3176
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 927 a 654 c 659 g 936 t
ORIGIN

Query Match 8.28; Score 73.4; DB 5; Length 3176;
Best Local Similarity 54.7%; Pred. No.3.9e-07;
Matches 146; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

OY 309 gtcaatagtagtaactctacaaaacgaagattcaaggaactatacttcgtgagttgaa 368
Db 1265 GACTTACTTAATCTGTGTCAGCAACATATCAAAAGTCCAAATCCGGTTAGCTATCG 1324
OY 369 tctgaagagcttaactcagttgtagtctgtacacacacatctcacgysaaatcccatc 428
Db 1325 TATCGTAACTTAAATACATGATGATCTTCACACACACAGATAAATGAGATCATTCCTTC 1384
OY 429 ttcttgggaaaattgaagtcactgttttttgcgcttaacgaaaacgattgacgg 488
Db 1385 TTCCCTTGATTTGGAGCATCTTTCACAGATGAACTTGAAGATATCAATACTG 1444
OY 489 tccatttccatagagaactcacagttattcaagccttaagttgtagtctcagysaa 548
Db 1445 TGTAACTTCCAGGCGACTTTGGAATCTAAGAGCATCATGGAATAGATCTTCAATGA 1504
OY 549 tgattgtggaacaattcacagtaga 575
Db 1505 TGAATCTCTGGCCCAATTCAGAGA 1531

Search completed: June 24, 2000, 00:23:23
Job time: 46358 sec

7
.
.
.

.

.....

.

.

1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:53:17 ; Search time 446.21 Seconds

(without alignments)
501.270 Million cell updates/sec

Title: US-09-180-798-28

Perfect score: 894
Sequence: 1 ggaacgattcaagcctccga.....gtcaattgattcttattata 894Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	894	100.0	894	1 V06589	Arabidopsis thalia
2	361	40.4	788	1 V06588	Arabidopsis thalia
3	361	40.4	1063	1 V06587	Arabidopsis thalia
4	356.2	39.8	981	1 V06587	Arabidopsis thalia
5	330	36.9	1106	1 V06586	Arabidopsis thalia
6	106	11.9	2089	1 V06591	Arabidopsis thalia
7	95.4	10.7	1814	1 V06571	Daucus carota SERK
8	73.4	8.2	3176	1 T62124	Arabidopsis thalia
9	71.6	8.0	4081	1 V06585	Arabidopsis thalia
10	66.4	7.4	3573	1 T06307	Partial tomato pat
11	66.4	7.4	6471	1 T06307	Tomato pathogen re
12	63	7.0	3541	1 V14522	CF-5 pathogen rest
13	63	7.0	6695	1 V06570	Daucus carota SERK
14	62.8	7.0	3979	1 V14518	CF-5 pathogen rest
15	62.8	7.0	3979	1 V14519	CF-5 pathogen rest
16	62.8	7.0	4123	1 V14523	CF-5 pathogen rest
17	62	6.9	1554	1 T31307	Tomato RRK gene cl
18	62	6.9	3293	1 X23532	Tomato Xa21 clone
19	61.2	6.8	3921	1 T31300	Rice Xa21 disease
20	61.2	6.8	13340	1 X23522	O. longistaminata
21	61.2	6.8	19639	1 X23524	O. longistaminata
22	59.6	6.7	7204	1 T31299	Rice Xa21 disease
23	59.6	6.7	7204	1 X23527	O. longistaminata X
24	54.8	6.1	8416	1 X23523	O. longistaminata
25	54.2	6.1	9424	1 X23525	O. sativa Xa21 gen
26	53.8	6.0	2880	1 O91450	Tomato Cf-9 CDNA
27	53.8	6.0	2880	1 T06304	Tomato pathogen re
28	53.8	6.0	3505	1 O91449	Tomato Cf-9 gene
29	53.8	6.0	3505	1 T06303	Tomato pathogen re
30	52.6	5.9	2192	1 X23530	Maize Xa21 gene DT
31	52.6	5.9	5940	1 X23526	O. longistaminata X
32	52.4	5.8	5733	1 X00477	Arabidopsis thalia
33	52.2	5.8	3050	1 T06309	Tomato pathogen re
34	52.2	5.8	3089	1 T47877	Tomato pathogen re

35 48.8 5.5 3045 1 X23531 Maize Xa21 gene DT
36 48.4 5.4 4104 1 X07356 Arabidopsis sterol
37 45.2 5.1 3842 1 X23533 Tomato Xa21 clone
38 40.6 4.5 2075 1 T49435 Tomato polygalactu
39 38.8 4.3 9295 1 T62125 Arabidopsis thalia
40 36.6 4.1 5361 1 T78868 P. falciparum live
41 36.6 4.1 6152 1 T78867 P. falciparum live
42 36 4.0 1117 1 Q42595 Sequence encoding
43 36 4.0 1140 1 V26098 Mandarin DNA regio
44 36 4.0 110000 1 V21209_15 Continuation (16 o
45 35.6 4.0 1380 1 V30822 Cytosolic glycerol

ALIGNMENTS

```
RESULT 1
ID V06589 standard; cDNA to mRNA; 894 BP.
AC V06589:
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS 1..678
FT FT /tag= a
FT FT /note= "shows high homology to SERK"
PN WO9743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
PI WPI; 98-086529/08.
DR P-PSDB; W47021.
DR Production of apomictic seeds - useful in plant breeding
PT Claim 28; Pages 83-84; 123pp; English.
PS The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences. 176 G; 285 T;
SQ Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;

Query Match 100.0%; Score 894; DB 1; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.4e-231;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggaacgattcaagcctccgaagggaagccttcaacgcgttgcgcgaactatcaagat 60
DB 1 GGAACGATTCAAGCCTCCGAAGGGAGCGCTTCAACGCGCTCCCGACCTTATCAGAT 60
QY 61 cgaacaatgtttgttcaagagtgtggatccaactctgttaacctgttactgtttcat 120
DB 61 CCAAGCAATGTTTGTTCAGAGTGGATCCCACTCTTGTATCTCTTACTGTGTTTCAT 120
QY 121 gtcacttgtaacacacacacacacacacacacacacacacacacacacacacac 180
DB 121 GTCACCTGTAACACACACACACACACACACACACACACACACACACACACACAC 180
QY 181 ggaacatctgtacccgaacttgggaagcttgaacattacacatactgtatggaatc 240
DB 181 GGAACATCTGTACCCGAACTTGGGAAGCTTGAACATTACACATATCTGTATGGAATC 240
QY 241 actctcttgccttcttgattatctgaacacattcatatataagtcacacataacatt 300
DB 241 ACTCTCTTGCCTTCTTGATTATCTGAAACATTTCATTATCAGACACATATTAACATT 300
QY 301 tgcattgaatcatatagtgaaactctacaaacagagattcaaggaactatcaccttga 360
DB 301 TGCATTGAATCATATAGTGAAGTCTACAAACAGATTCAGAACTATTAACCTTCGAG 360
QY 361 ctgtgaacatctgaagagctcaatcaagtttgatctgtacaacaacacacacacac 420
```

```
|||||
Db 361 CTTGAAATCTGAAGAGCTAATCAGTTGGATCTGTACAAACAATCTCACGGGAAA 420
Qy 421 atccattctcttgggaaatgtaagtcacttgttttttggggttaagaaacga 480
Db 421 ATCCCATCTCTTCTTGGGAAAATGTAAGTCACCTGTTTTTGGGCTTAAGAAACCGA 480
Qy 481 ttgacgcgtctctatccctagagaaactcacagttatctcaagccttaagttgttgaatc 540
Db 481 TTGACCGGTCCTATCTCTAGAGAACTCACAGTATTTCACGCTTAAGTTGTGATGTC 540
Qy 541 tcaaggaaatgatttctgtggaacaattccagttagaaggaaccttttgaacaattctcatg 600
Db 541 TCAGGGAAGATTTGTGTGAACAATCCAGTAGAAGACCTTTTGAACAATCCATG 600
Qy 601 caaaacttgagaaacaacttgattggagggagccagactactagttcttgcagctat 660
Db 601 CAAACTTTGAGAACACTGTGATTTGAGAGGACCAAGACCTAGTGTGCTGCAAGTAT 660
Qy 661 gacaccaattgcacttaaaagaagttgaagaacctataaagaagaatgttagtgacct 720
Db 661 GACACCAATTGCACCTAAAGAAAGTTGAGAACCTATTAAGAAAGATTTAGGTGACCT 720
Qy 721 ttagaagaactctgtaccagttgttgaatactatagagcctgttctatgtatata 780
Db 721 TGTAAAGACTGTGTACCAAGTGTGTAAATATATATAGAGCTGTGTTCATGTTATATA 780
Qy 781 tgaagcttggagagacagtaacttgcactgtatgttattgttgaagaaagttgaaatg 840
Db 781 TGAAGACTTTGAGAGACAGTAATCTGCATGTATTGTTGTTGTAAGAAATTTGAATG 840
Qy 841 agaatgtcttgaattgaattgtgttcttctatgttaacttgaattcttctta 894
Db 841 AGAATTGCTTGTAAATGGAATTTGTGTTCTTATGTAACCTGAATTTCTTATTA 894

RESULT 2
V06588 standard; cDNA to mRNA; 788 BP.
ID V06588;
AC 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomictic; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT 2..664
FT CDS
FT /tag= a
FT /note= "shows high homology to SERK"
MO9743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI; 98-086529/08.
DR P-PSDB; W47020.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 79-80; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;

Query Match 40.4%; Score 361; DB 1; Length 788;
Best Local Similarity 72.8%; Pred. NO. 3.7e-88;
Matches 536; Conservative 0; Mismatches 110; Indels 91; Gaps 2;
```

```
|||||
Db 151 CCAGAGCTTGAGTCCAAAGCTTGTAAATCTGTACCTGTTCCATGTCACCTTAACA 210
Qy 135 aacccatcaagtcactcgtctgatttgggaattcaactatctgacatcagtaacc 194
Db 211 AGACAAACCGCTCCTACTCGTGTGATTTGGGAAATCAACCTCTGTGACATCTTGGGCC 270
Qy 195 tgaacttggaaagcttgaacattcaaatatctgatatgaaatcactactcttgccttc 254
Db 271 TGAGTGTGGGAAGCTGAACATTATACAGTATCT----- 304
Qy 255 tgattctcgaagaaacttaccattatcagtcacacataataacatttgcgtacata 314
Db 304 ----- 304
Qy 315 tagtgactctcaaaaaagagattcaaggaactaactctcgtgacttgaacatctgaa 374
Db 304 ---AGAGCTCTACAAAACAAACATCCAAAGACTATACCTTCGAATCTGAAATCTGAA 360
Qy 375 gagcttaatcagtttgcgtctgatacacaacaattccacgggaaatccatctctt 434
Db 361 GAATCTCATCAGCTTGATCTGTACAAACAATCTTACAGGATAGTTCCACTCTTT 420
Qy 435 gggaaaattgaagtcactgttcttcttgggcttaacgaaaacgacttgcagctat 494
Db 421 GGGAAATTCAGTCTCTGCTCTTTTACGCTTATATGCAACGATTTGAGGGGCCAAT 480
Qy 495 tccatagagaactcacagttattcaagccttaagttgttgcgtcgaagaaatgattc 554
Db 481 CCTAGAGGCTCATCGCAATCCAAAGCCTTAAAGTTGTGTATGTCCTAAGCAATGATTT 540
Qy 555 gtctgaaacaattccagttgaaggaacctttgaacaactctctatgcaaaacttggaa 614
Db 541 GTGTGGAAACATCCAAACAAAGGACCTTTGCTCATCTCTTTACGAATTTGAGAA 600
Qy 615 caactgagatttgaaggaacagactactagttcttgcagctatgacccaattgac 674
Db 601 CAACCCGAGGTGAGAGGACCGGAATTAATCTCGTCTTGCACGTAAGCACTAATGAC 660
Qy 675 ttaaaagaagttgaaagacacttaagaagaa-tgttgggagacttgaagaaactgt 733
Db 661 CTGAAAAAATTTGGCAAAACCTGAAATTAAGAAATTTGGGGGTGACCTGTATAAGAACTT 720
Qy 734 taccagttgttgaact 752
Db 721 CACCACTTTATCAATATC 739

RESULT 3
V06590 standard; cDNA to mRNA; 1063 BP.
ID V06590;
AC 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomictic; apomictic; seeds; production; embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT 106..762
FT CDS
FT /tag= a
FT /note= "shows high homology to SERK"
MO9743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI; 98-086529/08.
DR P-PSDB; W47022.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 86-88; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
```

SQ Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;

Query Match

Best Local Similarity 40.4%; Score 361; DB 1; Length 1063;
Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2;

QY 15 ctccgaaggagcgccttcacgcgcttcgcggagctatcagatccagaaatgtgt 74
DB CTCCGAAGAGATGCTCTTACGCTTCGCGGAGTTAAAGATCCGACCATGTCT 248
QY 75 tcaagattggatccaactctgttaactccttgaactgtttcaatgacatgata 134
DB CCAAGGCTGGAGTCCACTTGTAACTGTAACTGGTTCATCTGACCTGTAACCA 308
QY 135 acacatcaagtcactcgtctgtgaatttggggaattcaaaattcttgaaatag 194
DB AGACAAACCGGCTCACTGCTGTGATTTGGGAATTCAAACCTCTGACATCTGGCC 368
QY 195 tgaacttggagcttgaaatttacaatatctgtatgaaatcaatcctttgtcctt 254
DB TGAGCTTGGAGGCTTGAACTTACAGTACT----- 402
QY 255 tgattatctgaaaacattacattacagtcacacataataaatttgccttgagcata 314
DB 402 ----- 402
QY 315 tagtgaactcaaaaaagagatccaagaaactataactcttgagcttggaaatgaa 374
DB 402 --AGAGCTTACAAAACAAACATCCAGGAACATATACCTCCGAACTTGGAAATCGAA 458
QY 375 gagctaatcagtttgaactctgtacaacaacatctcacccgggaaatcccatcttctt 434
DB 459 GAATCTCATACGCTTGATCTGTACAAACAATCTTACAGGGAATATCCCACTTCTT 518
QY 435 gggaaaattgaagtcacttgttttggcgcttaacgnaaacggatggacggtccat 494
DB 519 GGGAAAATTGAAGTCTGCTGCTTTTACGGCTTAATGACAAACGATTTGACGGGCGCAAT 578
QY 495 tccctagaagaactcacagatatttcaagccttaagtttgaattgtcacaagaaatgatt 554
DB 579 CCTTGAACACACTCATCTGCAATCCCAAGCTTAAAGTCTTGAATCTTCAAGCAATGATTT 638
QY 555 gtgtgaaacaattccagtagaagaagcctttgaaacaattccatgcaaaacttggaa 614
DB 639 GTGTGGAACAATCCCAAAACGAGACCTTGTGCTACATCTTACAGAACTTTGAGAA 698
QY 615 caacttgagattggaggaagcaagaactactagtccttggcgagctatgacacaaattgcac 674
DB 699 CAACCCGAGGTTGGAGGACCGGAATTTCTGCTTGAAGCTTACGACACTTACCTGAC 758
QY 675 ttaaaagaagttgaagaacctataaagaaga-tgttaggtgaccttgaagaactctg 733
DB 759 CTGAAAATAATTGCAAAACCTGAATAATGAGAAATTGGGGGTGACCTTGAAGAACACTT 818
QY 734 taccgaagtgttgaataac 752
DB 819 CACCACTTATCAAAATATC 837

RESULT 4

V06587 standard; cDNA to mRNA; 981 BP.

AC V06587;

DT 03-ADG-1998 (first entry)

DE Arabidopsis thaliana SERK LRR homologous EST clone.

KW receptor kinase; apomixis; apomictic; seeds; production; embryos;

OS Arabidopsis thaliana.

FT Key Location/Qualifiers

FT CDS 104..760

FT /tag= a

FT /note= "shows high homology to SERK"

PN W09743427-A1.
PD 20-NOV-1997.
PR 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI; 98-086529/08.
DR P-PSDB; W47019.
PS Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 75-77; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.

SQ Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 39.8%; Score 356.2; DB 1; Length 981;
Best Local Similarity 72.4%; Pred. No. 7.8e-87;
Matches 535; Conservative 0; Mismatches 113; Indels 91; Gaps 2;

QY 15 ctccgaaggagcgccttcacgcgcttcgcggagctatcagatccagaaatgtgt 74
DB 187 CTCCGAAGAGATGCTCTTACGCTTCGCGGAGTTGACAGATCCGACCATGTCT 246
QY 75 tcaagattggatccaactctgttaactccttgaactgtttcaatgacatgata 134
DB 247 CCAAGGCTGGAGTCCACTTGTAACTGTAACTGGTTCATCTGACCTGTAACCA 306
QY 135 acacatcaagtcactcgtctgtgaatttggggaattcaaaattcttgaaatag 194
DB 307 AGACAAACCGGCTCACTGCTGTGATTTGGGAATTCAAACCTCTGACATCTGGCC 366
QY 195 tgaacttggagcttgaaatttacaatatctgtatgaaatcaatcctttgtcctt 254
DB 367 TGAGCTTGGAGGCTTGAACTTACAGTACT----- 400
QY 255 tgattatctgaaaacattacattacagtcacacataataaatttgccttgagcata 314
DB 400 ----- 400
QY 315 tagtgaactcaaaaaagagatccaagaaactataactcttgagcttggaaatgaa 374
DB 400 --AGAGCTTACAAAACAAACATCCAGGAACATATACCTCCGAACTTGGAAATCGAA 456
QY 375 gagctaatcagtttgaactctgtacaacaacatctcacccgggaaatcccatcttctt 434
DB 457 GAATCTCATACGCTTGATCTGTACAAACAATCTTACAGGGAATATCCCACTTCTT 516
QY 435 gggaaaattgaagtcacttgttttggcgcttaacgnaaacggatggacggtccat 494
DB 517 GGGAAAATTGAAGTCTGCTGCTTTTACGGCTTAATGACAAACGATTTGACGGGCGCAAT 576
QY 495 tccctagaagaactcacagatatttcaagccttaagtttgaattgtcacaagaaatgatt 554
DB 577 CCTTGAACACACTCATCTGCAATCCCAAGCTTAAAGTCTTGAATCTTCAAGCAATGATTT 636
QY 555 gtgtgaaacaattccagtagaagaagcctttgaaacaattccatgcaaaacttggaa 614
DB 637 GTGTGGAACAATCCCAAAACGAGACCTTGTGCTACATCTTACAGAACTTTGAGAA 696
QY 615 caacttgagattggaggaagcaagaactactagtccttggcgagctatgacacaaattgcac 674
DB 697 CAACCCGAGGATTGGAGGACCGGAATTTACTGCTTGAAGCTTACGACACTTACCTGAC 756
QY 675 ttaaaagaagttgaagaacctataaagaaga-tgttaggtgaccttgaagaactctg 733
DB 757 CTGAAAACAACTGGCAAAACCTGAATAATGAGAAATTGGGGGTGACCTTGAAGAACACTT 816
QY 734 taccgaagtgttgaataac 752
DB 817 CACCACTTATCAAAATATC 835

RESULT 5

Db	459	CTTGGTGCTCTCAAAATTTCCAGATTGTGA-----CCTTACAGT	5001
Qy	259	tattcgaanaacattacattatagtcacacataaacttttgcttggatcatagt	318
Db	501	AACAACATACACGCGCCGCAATTCATGTAATCTGGAAATCTGACAACTATGTAAGTTTG	5605
Qy	319	gaacctacacaacagagatccaagaaactaacctctcgaactctggaatctgaagagt	3789
Db	561	GACTTTACTTAAACACGCTTCGCGGTCTATTCCGGAATCATTTGGGAAAGCTTTCAAG	6200
Qy	379	ctaatcagttggatctctgcaacaacaatctaccggaaatcccatctctttggga	4388
Db	621	CTGAGATTTCTCCGGTTAAACAACAACAGTCTACGTGCGGCAATTCATATGCACTGACC	6800
Qy	439	aaattgaagtactcgttttttttgcgcttaagaaaaaccgatctgaccggtccatctct	4588
Db	681	AATATTAATCTACCTTCAAGTGTAGATATATCAATATACAGACTCTCTGTCACTTCT	7400

RESULT	7
ID	V06571
AC	V06571 standard; cDNA; 1814 BP.
AD	V06571;
DT	03-AUG-1998 (first entry)
DE	Daucus carota SERK gene.
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;
KM	plant breeding; ss.
OS	Daucus carota.
FT	key
FT	location/qualifiers
FT	94..1755
FT	/*tag= a
FT	/product= SERK protein
PN	MO9743427-A1.
PD	20-NOV-1997.
PE	13-MAY-1997; E02443.
PR	14-MAY-1996; GB-010044.
PA	(NOVS) NOVARTIS AG.
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR	WPI: 98-086529/08.
DR	P-P5DB: W47013.
PT	Production of apomictic seeds - useful in plant breeding
PS	Claim 21: Pages 47-51; 123pp; English.
CC	The sequence is that encoding SERK, a putative receptor kinase.
CC	It may be used as part of a method of producing apomictic seeds
CC	comprising: (a) transforming plant material with a nucleotide
CC	sequence encoding a protein which in active form in a cell or
CC	cell membrane renders the cell embryogenic; (b) regenerating
CC	the transformed material into plants or carpel-containing
CC	plant parts; and (c) expressing the sequence in the vicinity
CC	of the embryo sac. The apomictic seeds and embryos thus produced
CC	can be developed into plant progeny. This is useful in plant
CC	breeding programs. Controllable and reproducible apomixis provides
CC	many advantages in plant improvement and cultivar development in
CC	the case that sexual plants are available as crosses with the
CC	apomictic plant. Apomixis provides for true-breeding, seed
CC	propagated hybrids and could shorten and simplify the breeding
CC	process so that selfing and progeny testing to produce and/or
CC	stabilise a desirable gene combination could be eliminated.
CC	Apomixis allows plant breeders to develop cultivars with
CC	specific stable traits for such characteristics as height,
CC	seed and forage quality and maturity.
CC	Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

Query Match	10.7%;	Score 95.4;	DB 1;	Length 1814;
Best Local Similarity	56.1%;	Pred. No. 1.5e-16;		
Matches 180;	Conservative	0;	Mismatches 141;	Indels 0;
				Gaps 0;

```

Oy 319 gaactctacaaaacgagattcaaggaactatacccttcgtgagcttgyaaactcgaagagt 378
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 178 GAGCTTACAGCATAACATAAGTGAGACCAATCTCTAGTCATCTTGGGAATCTGACAAAT 237

```

OY	379	ctaatcagtttgatctgtaacacacatctcacccggaaatcccatctctttggga	438
Db	238	TTGTGAGCTTGGACCTATACATGAATAGCTCTCTGGACCTATACGGACAAATTAGA	297
OY	439	aaattgaagtcacttggttttttttcggccttaagaacaccgattacgcggtccattcc	498
Db	298	AAGCTTACAAAGGCTAAGATCTCTGGCTCACAACAAAGCCTCTCTGGTCAATTGCA	357
OY	499	agagaactcacgattatccaagccttaagttgtagtctccaggaaatgatttggt	558
Db	358	ATGTCACTGACTAATTATTCAACTCTTAACTCCTGGATTATCAACAATGGCTATCA	411
OY	559	ggaacaattccagtagaaggaccttttgaacacattcctatgcaaaactttagaacaac	611
Db	418	GGACCACTAACCGGATATGGCTCATTTTCTTTTACACCTATCAGTTTTCACATATAT	477
OY	619	cttagagattgaggagacagaa	639
Db	478	TTGAATTATGTGACCCGTA	498

```

RESULT      8
ID          T62124
AC          T62124; standard; cDNA to mRNA; 3176 BP.
DE          10-JUN-1997 (first entry)
DI          Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
KW          Plant; morphogenesis; regulation; short; stem; alteration;
KW          inflorescence; extraneous; gene; expression; transformation;
KW          increase; control; form; length; ds.
OS          Arabidopsis thaliana.
FH          Key
FT          Location/Qualifiers
FT          cds
FT          51..2981
FT          /*tag=a
FT          /note="Plant morphogenesis regulatory protein"
PN          J09056382-A.
PD          04-MAR-1997.
PE          24-AUG-1995; 216187.
PF          24-AUG-1995; JF-216187.
PA          (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA          (CHIK-) ZH CHIKU KANKYO SANGYO GUTSU KENKYU.
DR          WPI: 97-206629/19.
DR          P-BSDB: M13408.
PT          DNA encoding plant morphogenesis regulatory protein - useful to
PT          yield plants with short stems or altered inflorescence
PS          Claim 1; Pages 6-10; 17pp; Japanese.
CC          The present sequence encodes an Arabidopsis thaliana plant
CC          morphogenesis regulatory protein (MRP), which can be used to yield
CC          a plant with, e.g. short stems or altered inflorescence. The MRP
CC          acts on a plant at a specific site for a specific period, and can
CC          therefore be used to regulate extraneous gene expression in a
CC          plant. The MRP's cDNA or genomic DNA can be used to transform a
CC          plant to increase its MRP expression, and therefore control the
CC          form (particularly stem length) of the plant.
SQ          Sequence 3176 BP: 927 A; 654 C; 936 T;

```

Query Match	8.28;	Score 73.4;	DB 1;	Length 3176;
Best Local Similarity	54.78;	Pred. No. 1.6e-10;		
Matches 146; Conservative	0;	Mismatches 121;	Indels 0;	Gaps 0

DY 309 gtcattatgtaacctctacaacaaacggagattcaaggacatacctctgagcttggaaa 368
| | | | | | | | | | | | | | |
Db 1265 GACTTACCTTATCTGTCCAGCAACAATTATCAAGAATCCCGGTTGAGCATCTCG 1324

DQ 369 tctgaagatcctaactcaagtcttgatctgtacacaacaactctcacggyaaaaaccatc 428
| | | | | | | | | | | | | | | | | |
Db 1325 TATCGGTACTTAGTACATCGATGTCCTTTCCACAACAAGATAAATGGATCATTTCCCTTC 1384

Dy 429 tctttggyaaaattgaagtcacttgttttttgcgcgttaacgaaaaocgatggaccgg 488
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 TTCCCTTGGTGATTGGAGCACTTTCACAGATGAAC'TGAGTTAGAATCATATAACTCG 1444

Oy		489	tccattcttcagagaacatcacagtatttcaaggcttaaggttgtagtccagggaa	548
Db	1445	TGTAATTCACAGGCACATTGGAAATCTAAGAAGCATCATGGAATAGACTTTCAATAA	150	
Oy		549	tgatttgtgtggaacaattccagtaga	575
Db	1505	TGATATCTGTGCCCAATTCAGAGA	1531	
RESULT	9			
V06585				
ID	V06585	standard; DNA; 4081 BP.		
AC	V06585;			
DT	03-AUG-1998	(first entry)		
DE	Arabidopsis thaliana SERK gene.			
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;			
KM	plant breeding; ds.			
OS	Arabidopsis thaliana.			
FX	Key	Location/Qualifiers		
FT	CDS	3696..6620		
FT		/tag= a		
FT	Intron	/note= "contains introns"		
FT		3731..3802		
FT		/tag= b		
FT	Intron	/number= 1		
FT		3851..3979		
FT		/tag= c		
FT	Intron	/number= 2		
FT		4124..4211		
FT		/tag= d		
FT	Intron	/number= 3		
FT		4284..4357		
FT		/tag= e		
FT	Intron	/number= 4		
FT		4430..4528		
FT		/tag= f		
FT	Intron	/number= 5		
FT		4642..4757		
FT		/tag= g		
FT	Intron	/number= 6		
FT		4890..4967		
FT		/tag= h		
FT	Intron	/number= 7		
FT		5295..5803		
FT		/tag= i		
FT	Intron	/number= 8		
FT		6197..6339		
FT		/tag= j		
FT		/number= 9		
PN	WO9743427-A1.			
PD	20-NOV-1997.			
PF	13-MAY-1997; E02443.			
PI	14-MAY-1996; GB-010044.			
PA	(NOVS) NOVARTIS AG.			
PR	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;			
DR	MP1: 98-086529/08.			
DR	P-PsDB; W47017.			
PT	Production of apomictic seeds - useful in plant breeding			
PS	Claim 26; Pages 64-67; 123pp; English.			
CC	The sequence is that encoding SERK, a putative receptor kinase.			
CC	It may be used as part of a method of producing apomictic seeds			
CC	comprising: (a) transforming plant material with a nucleotide			
CC	sequence encoding a protein which in active form in a cell or			
CC	cell membrane renders the cell embryogenic; (b) regenerating			
CC	the transformed material into plants or carpel-containing			
CC	plant parts; and (c) expressing the sequence in the vicinity			
CC	of the embryo sac. The apomictic seeds and embryos thus produced			
CC	can be developed into plant progeny. This is useful in plant			
CC	breeding programs. Controllable and reproducible apomixis provides			
CC	many advantages in plant improvement and cultivar development In			
CC	the case that sexual plants are available as crosses with the			
CC	apomictic plant. Apomixis provides for true-breeding, seed			
CC	propagated hybrids and could shorten and simplify the breeding			

CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 4081 Bp; 1120 A; 770 C; 785 G; 1406 T;

Query Match	8.0%;	Score 71.6;	DB 1;	Length 4081;
Best Local Similarity	54.4%;	Pred. No. 5.2e-10;		
Matches 186;	Conservative	0;	Mismatches 129;	Indels 27;
				Gaps 1

OY	157	gatttgggaatccaacttatcttgagacccagtaacctgcgaacttgggaacttcgaaact	216
Dd	2015	GATTTGGGAATCGAGGTTATCTGCSCAATTGTCCAGACCTTGGTGTCCTCAAGAA	2074
OY	217	ttacaatatcgtatgy-----atacataaccctttg	249
Dd	2075	TTCGAGATTATTGTAAGTCCACTTATGATCATNGCTTTAACAAAACAATAATCCAGATTG	2134
OY	250	ccttgtattatctcgaaaacattacattatcaagtacacatatcacatttgccttag	309
Dd	2135	ACGAAGAAGACCTCGAGTTACCTTTTGTGAATTAAGAACTTTTAAACAAGTTCTTATT	2194
OY	310	tcatatagtgaactctacaaaaacgagatcaaaggaaactatccccttcgagcttggaaat	369
Dd	2195	TCTTACAGGGAGCTTTACATTAACAACATPAACTGGGCCGATTCCTAGTAATCTTGGAAAT	2254
OY	370	ctgaagaagctcaatcattgatctgtcaacaacaactccatccoggggaataatccccact	429
Dd	2255	CTGCAAACTTAGTAGATTTTGGATCTTTACTTPAAACAGCTTCTCCGGGTCCTATTCCGGA	2314
OY	430	tctttgggaaaattgaaagtcaacttgttttttgcggttaac	471
Dd	2315	TCATTGGGAAAGCTTTCAAAGCTGAGAAATCTCTGTGATPAC	2356

RESULT 10

106307

ID 106307 standard; cDNA; 3573 bp.

AC 106307;

DE 14-APR-1996 (first entry)

DE Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.

KN Pathogen resistant; Cf-2.2; tomato; C.fulvum; AVR 4; AVR 9; fungal;

KW leaf mould; variegation; ss.

OS Lycopersicon esculentum.

PN W05531564-A2.

PD 23-NOV-1995.

PE 11-MAY-1995; G01075.

PR 11-MAY-1994; GB-009394.

PR 23-DEC-1994; WO-G02812.

PR 31-MAR-1995; GB-006658.

PR 07-APR-1995; GB-007232.

PA (GARS-) GATSBY CHARITABLE FOUND.

PI Hammond-Kosack KE, Jones DA, Jones JDC;

DR P-PSDB; R85299.

PT Increasing plant pathogen resistance by induction of variegation

PT may lead to acquired resistance to a broad range of pathogens.

PS Claim 9; Page 85-87; 13pp; English.

CC 106307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone

CC in a new method this gene is expressed highly in genetic constructs

CC which may be used to impart a broad range of pathogen resistance, by

CC induction of variegation, to transgenic plants (or parts or propagules

CC of plants) containing such constructs. Cf-2.2 imparts resistance to the

CC disease caused by the leaf mould fungal pathogen Cladosporium fulvum.

CC C.fulvum contains avirulence (Avr) genes that confer recognition by

CC plants containing Cf genes, leading to the activation of host defence

CC mechanisms to attack the disease.

CC Sequence 3573 bp; 1032 A; 654 C; 664 G; 1223 T;

7.48; Score 66.4; DB 1; Length 3573;

Best Local Similarity 54.0%; Pred. No. 1.2e-08;
Matches 136; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
QY 320 aactcacaacagagattcaaggaactataccttcgagcttggaatctgaagatc 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1230 ATCTCTACATATATACAGCTTTCGCTCTATTCCTGAATAATAGTACTTGACTCTC 1289
QY 380 taatcagtttgatctgtacaacaacatctcccggaagaatcccatctctctggaa 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1290 TTAATTATCTAGATTGAGTAATACCTCATTAATGATTTATTCCTGCTCATTTGGCA 1349
QY 440 aattgaagtaactgttttttcgagcttaagaaacgagatgacggtctctactcta 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1350 ATATGACCAACTGGCTTTTGTCTTTTATGAAATACGCTTCTGCTCTGTTCTCG 1409
QY 500 gagaactcaacagttattcaagccttaagttgtatgtctcagggaatgtttgttg 559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1410 AAGAAATAGGTTACCTAAGGTCCTTAATGCTCTGATTGAGTGAATGCTCTTAATG 1469
QY 560 gaacaattccag 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1470 GCTCTATTCTCG 1481
```

RESULT 11

```
T06306
ID T06306 standard; DNA; 6471 BP.
AC T06306;
DT 14-APR-1996 (first entry)
DE Tomato pathogen resistance gene Cf-2.1.
KW Pathogen resistant; Cf-2.1; tomato; C.fulvum; Avr 4; Avr 9; fungal;
KM leaf mould; variegation; ds.
OS Lycopersicon esculentum.
FH Key location/Qualifiers
FT 5'utr 1..1676
FT cds /*tag- a
    1677..5014
    /*tag- b
    /*product- immature_Cf-2.1_protein
FT signal_peptide /*tag- c
    1677..1745
    /*tag- d
    1746..5011
    /*tag- e
    5015..6471
    /*tag- e
FT 3'utr
FT
FT W09531564-A2.
PN 23-NOV-1995.
PF 11-MAY-1995; G01075.
PR 11-MAY-1994; GB-008394.
PR 23-DEC-1994; WO-G02812.
PR 31-MAR-1995; GB-006658.
PR 07-APR-1995; GB-007232.
PA (GATS-) GATSBY CHARITABLE FOUND.
PI Hammond-Kosack KE, Jones DA, Jones JDG;
DR WPI: 96-010949/O1.
DR P-PSDB; R85298.
FT Increasing plant pathogen resistance by induction of variegation -
    PT may lead to acquired resistance to a broad range of pathogens.
PS Claim 9; Page 80-83; 131pp; English.
CC T06306 is the tomato pathogen resistance gene Cf-2.1. In a new method
    CC this gene is expressed highly in genetic constructs which may be used
    CC to impart a broad range of pathogen resistance, by induction of
    CC variegation, to transgenic plants (or parts or propagules of plants)
    CC containing such constructs. Cf-2.1 imparts resistance to the disease
    CC caused by the leaf mould fungal pathogen Cladosporium fulvum.
    CC C.fulvum contains avirulence (Avr) genes that confer recognition by
    CC plants containing Cf-genes, leading to the activation of host
    CC defence mechanisms to attack the disease.
SQ Sequence 6471 BP; 2073 A; 1106 C; 1122 G; 2170 T;
```

Query Match 7.4%; Score 66.4; DB 1; Length 6471;
Best Local Similarity 54.0%; Pred. No. 1.6e-08;

Matches 136; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
QY 320 aactcacaacagagattcaaggaactataccttcgagcttggaatctgaagatc 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2914 ATCTCTACATATATACAGCTTTCGCTCTATTCCTGAATAATAGTACTTGACTCTC 2973
QY 380 taatcagtttgatctgtacaacaacatctcccggaagaatcccatctctctggaa 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2974 TTAATTATCTAGATTGAGTAATACCTCATTAATGATTTATTCCTGCTCATTTGGCA 3033
QY 440 aattgaagtaactgttttttcgagcttaagaaacgagatgacggtctctactcta 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3034 ATATGACCAACTGGCTTTTGTCTTTTATGAAATACGCTTCTGCTCTGTTCTCG 3093
QY 500 gagaactcaacagttattcaagccttaagttgtatgtctcagggaatgtttgttg 559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3094 AAGAAATAGGTTACCTAAGGTCCTTAATGCTCTGATTGAGTGAATGCTCTTAATG 3153
QY 560 gaacaattccag 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3154 GCTCTATTCTCG 3165
```

RESULT 12

```
V14522
ID V14522 standard; DNA; 3541 BP.
AC V14522;
DT 20-MAY-1998 (first entry)
DE Cf-5 pathogen resistance gene clone Hcr2-5b gene.
KW Tomato; Cf-5 pathogen resistance gene; Cladosporium fulvum-5;
KM tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon pimpinellifolium.
FH Key location/Qualifiers
FT CDS 603..3002
FT /*tag- a
    /*product- Cf-5 pathogen resistance gene clone Hcr2-5B
FT PN W09743429-A1.
FT 20-NOV-1997.
PF 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1996; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR WPI: 98-008895/O1.
DR P-PSDB; W41311.
FT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
    PT for production of transgenic plants resistant to pathogens e.g.
    PT tomato leaf mould C.fulvum in tomatoes
PS Disclosure; Fig 5; 75pp; English.
CC This sequence is a clone of the polynucleotide of the invention, which
    CC is able to confer pathogen resistance on a plant. The polynucleotide of
    CC the invention is one of two tomato Cf-5 gene variants, which offer
    CC resistance against the pathogen Cladosporium fulvum-5. Transgenic plants
    CC can be produced by incorporating the gene into plant cells and
    CC regenerating plants from the cells; asexually or sexually produced
    CC offspring can also be subsequently produced. Expression of the gene in
    CC plant cells can confer pathogen resistance on a plant e.g. to tomato leaf
    CC mould (C. fulvum) in tomatoes. Oligonucleotides with sequences
    CC complementary to the gene or fragments of it, are useful in anti-sense
    CC techniques to reduce gene expression. The nucleic acids/polynucleotides
    CC are useful as hybridisation probes to identify other genes/fragments
    CC conferring pathogen resistance on plants e.g. Phytophthora resistance in
    CC potatoes. Homologues between Cf-5 and Cf-9 may be used to identify
    CC further resistance genes of this class.
SQ Sequence 3541 BP; 1123 A; 630 C; 612 G; 1176 T;
```

Query Match 7.0%; Score 63; DB 1; Length 3541;
Best Local Similarity 52.5%; Pred. No. 1e-07;
Matches 138; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 309 gtcatatagtgactctacacaacagattcaaggaactataccttcgagcttggaagaa 368
||| | || | |||| | | | | || | |||| | | | |

Query Match 7.0%; Score 62.8; DB 1; Length 3979;
 Best Local Similarity 53.2%; Pred. No. 1.2e-07;
 Matches 133; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

QY 320 aactctacaaaagagattcaaggaactatatacctcttgagcttggaaatctgaagatc 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2107 ATCTTTACAAATATACAGCTTTCGCTCTATTCCTGCTTATTTGGCAATATAGAAATC 2166
QY 380 taatcaagttggatcgtgtacacaacatctcacccgggaaatccatctcttggaa 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2167 TGCAAACTCTGTTTCACGATTAACGATCTCATTTGGGAAATTCCTCATTTGTGTGCA 2226
QY 440 aattgaagtcactgttttttgcggttaacgaaacccgaltgacccggtcctatccta 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2227 ATTTGACATCACTGGAAGTGTGTATATGTCGAGAAACAATTTGAAGGAAAGTTCCGC 2286
QY 500 gagaaactcacagttattcaagccttaagttgtagtctcagggaaatgattgtgtg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2287 AATGTTGGTAATATATAGTACCTTCACATTTTGTGATGTATATATAGTTTCAAGAG 2346
QY 560 gaacaattcc 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2347 GAGAGCTCCC 2356
  
```

RESULT 15

```

V14519
ID V14519 standard; DNA; 3979 BP.
AC V14519;
DE 20-MAY-1998 (first entry)
DE CF-5 pathogen resistance gene variant #2.
KW Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon pimpinellifolium.
FH Key Location/Qualifiers
FT CDS 653..3560
    /*tag= a
    /product= CF-5 pathogen resistance gene
PN MO9743429-A1.
PD 20-NOV-1997.
PE 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1996; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR P-PEDB: M41310.
PT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
    for production of transgenic plants resistant to pathogens e.g.
    tomato leaf mould C. fulvum in tomatoes
PR Claim 6; Fig 1b; 75pp; English.
CC This sequence is an example of the polynucleotide of the invention, and
    is able to confer pathogen resistance on a plant. It is one of two tomato
    CF-5 gene variants, which offer resistance against the pathogen
    Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
    the gene into plant cells and regenerating plants from the cells;
    asexually or sexually produced offspring can also be subsequently
    produced. Expression of the gene in plant cells can confer pathogen
    resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
    Oligonucleotides with sequences complementary to the gene or fragments
    of it, are useful in anti-sense techniques to reduce gene expression. The
    nucleic acids/polynucleotides are useful as hybridisation probes to
    identify other genes/fragments conferring pathogen resistance on plants
    e.g. Phytophthora resistance in potatoes. Homologies between Cf-5 and
    Cf-9 may be used to identify further resistance genes of this class.
SQ Sequence 3979 BP; 1218 A; 716 C; 698 G; 1347 T;
  
```

Query Match 7.0%; Score 62.8; DB 1; Length 3979;
 Best Local Similarity 53.2%; Pred. No. 1.2e-07;
 Matches 133; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

QY 320 aactctacaaaagagattcaaggaactatatacctcttgagcttggaaatctgaagatc 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2107 ATCTTTACAAATATACAGCTTTCGCTCTATTCCTGCTTATTTGGCAATATAGAAATC 2166
QY 380 taatcaagttggatcgtgtacacaacatctcacccgggaaatccatctcttggaa 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2167 TGCAAACTCTGTTTCACGATTAACGATCTCATTTGGGAAATTCCTCATTTGTGTGCA 2226
QY 440 aattgaagtcactgttttttgcggttaacgaaacccgaltgacccggtcctatccta 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2227 ATTTGACATCACTGGAAGTGTGTATATGTCGAGAAACAATTTGAAGGAAAGTTCCGC 2286
QY 500 gagaaactcacagttattcaagccttaagttgtagtctcagggaaatgattgtgtg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2287 AATGTTGGTAATATATAGTACCTTCACATTTTGTGATGTATATATAGTTTCAAGAG 2346
QY 560 gaacaattcc 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2347 GAGAGCTCCC 2356
  
```

Search completed: June 23, 2000, 22:53:35
 Job time: 40611 sec

1

2

3

4

5

6

7

8

9

10

11

12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:45:40 ; Search time 274.21 Seconds

(without alignments)
423.787 Million cell updates/sec

Title: US-09-180-798-28

Sequence: 1 ggaccgattcaagcctccga.....gtacttgattcttatia 894

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_MA:*
1: /cgn2_6/prodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/prodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/prodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/prodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/prodata/2/1na/6.COMB.seq:*
6: /cgn2_6/prodata/2/1na/PCRTUS.COMB.seq:*
7: /cgn2_6/prodata/2/1na/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	62	6.9	1554	4	US-08-587-680A-24
2	61.2	6.8	3921	4	US-08-587-375-3
3	61.2	6.8	3921	4	US-08-587-680A-3
4	61.2	6.8	3921	4	US-08-587-680A-3
5	59.6	6.7	6256	3	US-08-475-891A-3
6	59.6	6.7	6256	3	US-08-475-891A-1
7	59.6	6.7	6256	4	US-08-567-375-1
8	58.8	6.6	831	4	US-08-587-680A-1
9	58.8	6.6	831	4	US-08-567-375-15
10	53.8	6.0	2880	4	US-08-666-271-4
11	52.4	5.9	3905	4	US-08-666-271-1
12	40.6	4.5	3733	3	US-08-473-533A-1
13	36	4.0	1116	2	US-08-238-163-3
14	35.6	4.0	1380	5	US-08-244-646-14
15	35.4	4.0	3095	7	US-08-968-563-5
16	35	3.9	19124	4	US-08-487-826B-13
17	34.4	3.8	1058	1	US-08-238-163-1
18	34.4	3.8	2917	2	US-08-992-936B-20
19	34.4	3.8	2917	2	US-09-111-573-20
20	34.4	3.8	7218	1	US-08-232-463-14
21	33.8	3.8	1868	4	US-08-960-022-5
22	33.6	3.8	3440	2	US-08-471-791-27
23	33.6	3.8	3440	6	PCT-US81-01746-27
24	32.8	3.7	792	2	US-08-244-646-16
25	32.8	3.7	2210	1	US-07-710-361-2
26	32.2	3.6	4599	2	US-08-431-080-27
27	32.2	3.6	4599	4	US-08-938-534-27

28	32.2	3.6	4951	4	US-08-867-030B-5	Sequence 5, Appl1
29	32.2	3.6	4951	6	PCT-US95-06119-5	Sequence 5, Appl1
30	31.2	3.5	2241	6	PCT-US95-10509-1	Sequence 1, Appl1
31	31	3.5	426	2	US-08-442-063A-35	Sequence 35, Appl1
32	31	3.5	567	2	US-08-442-063A-38	Sequence 38, Appl1
33	31	3.5	711	2	US-08-442-063A-41	Sequence 41, Appl1
34	31	3.5	849	2	US-08-442-063A-44	Sequence 44, Appl1
35	31	3.5	909	2	US-08-347-471-3	Sequence 3, Appl1
36	31	3.5	924	2	US-08-442-063A-47	Sequence 47, Appl1
37	31	3.5	1002	2	US-08-442-063A-26	Sequence 26, Appl1
38	31	3.5	1026	2	US-08-272-919-1	Sequence 1, Appl1
39	31	3.5	1026	6	US-08-619-916-1	Sequence 1, Appl1
40	31	3.5	1026	6	PCT-US95-08542-1	Sequence 1, Appl1
41	31	3.5	1593	7	5340934-3	Patent No. 5340934
42	30.8	3.4	2674	4	US-08-926-724-2	Sequence 2, Appl1
43	30.6	3.4	2044	1	US-08-071-601-3	Sequence 3, Appl1
44	30.6	3.4	2044	7	US-08-621-100-3	Sequence 3, Appl1
45	30.6	3.4	11225	7	5182210-9	Patent No. 5182210

ALIGNMENTS

RESULT 1
US-08-587-680A-24
Sequence 24, Application US/08587680A
Patent No. 5977434
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587, 680A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373, 375
FILING DATE: 17-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475, 891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004, 645
FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/567, 375
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058940US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1554 base pairs

1 AFFILIATION NUMBER: US 06/515,513
2
3 FILING DATE: 17-JAN-1995

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/475,891
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/004,645
;; FILING DATE: 29-SEP-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/567,375
;; FILING DATE: 04-DEC-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bastian, Kevin L.
;; REGISTRATION NUMBER: 34,774
;; REFERENCE/DOCKET NUMBER: 023070-058940US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3921 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: join(1..2676, 3520..3918)
;; OTHER INFORMATION: /product= "Xa-21"
US-08-587-680A-3

Query Match      6.8%; Score 61.2; DB 4; Length 3921;
Best Local Similarity 52.3%; Pred. No. 1.1e-08;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 313 tataggaactcacaaaacgagatcacaagaactacattccttgagcttggaaatctg 372
DB 1138 TTCTTGCACTTGAAATGATAGATACAGAGACATTCGAGAGATTTGGCAATCTT 1197
QY 373 aagaagctaatcagcttgatctgtacacaacaatctcaccgggaaatcccatctct 432
DB 1198 ATTGCGTTACAACTCTATCTCTGCAACAACAATTTCAGAGGGCTCTTCCATCATCG 1257
QY 433 ttgggaattgaagtcactgttttttgcggttaacgaaacggattgacggctct 492
DB 1258 TTGGGAGGCTTAAAACTTAGGCATTTCTACGCGCTACGAAACAACCTTGAGCGGTTG 1317
QY 493 attccctagagaactcacagttattcaagccttaaggttgatgctcagggatgat 552
DB 1318 ATCCCGTTGGCCATAGGAATCTTACTGAATTAATATCTTACTGCTCGGACCAACAA 1377
QY 553 ttgtgtggaacaattcca 570
DB 1378 TTCAGTGTGTGATACCA 1395

RESULT 4
US-08-475-891A-3
; Sequence 3, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
```

```
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/475,891A
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/373,375
;; FILING DATE: 17-JAN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bastian, Kevin L.
;; REGISTRATION NUMBER: 34,774
;; REFERENCE/DOCKET NUMBER: 02370-058910US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5992 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: join(512..3149, 3993..4393)
;; OTHER INFORMATION: /product= "RRK-B"
;; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
;; OTHER INFORMATION: resistance gene RRK-B from Rice (Oryza
;; OTHER INFORMATION: sativa)"
US-08-475-891A-3

Query Match      6.8%; Score 61.2; DB 3; Length 5992;
Best Local Similarity 52.3%; Pred. No. 1.4e-08;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 313 tataggaactcacaaaacgagatcacaagaactacattccttgagcttggaaatctg 372
DB 1649 TTCTTGCACTTGAAATGATAGATACAGAGACATTCGAGAGATTTGGCAATCTT 1708
QY 373 aagaagctaatcagcttgatctgtacacaacaatctcaccgggaaatcccatctct 432
DB 1709 ATTGCGTTACAACTCTATCTCTGCAACAACAATTTCAGAGGGCTCTTCCATCATCG 1768
QY 433 ttgggaattgaagtcactgttttttgcggttaacgaaacggattgacggctct 492
DB 1769 TTGGGAGGCTTAAAACTTAGGCATTTCTACGCGCTACGAAACAACCTTGAGCGGTTG 1828
QY 493 attccctagagaactcacagttattcaagccttaaggttgatgctcagggatgat 552
DB 1829 ATCCCGTTGGCCATAGGAATCTTACTGAATTAATATCTTACTGCTCGGACCAACAA 1888
QY 553 ttgtgtggaacaattcca 570
DB 1889 TTCAGTGTGTGATACCA 1906

RESULT 5
US-08-475-891A-1
; Sequence 1, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,891A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 02370-05891005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1648..4383, 5178..5513)
OTHER INFORMATION: /product= "RRK-F"
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
OTHER INFORMATION: sativa)"
US-08-475-891A-1

Query Match 6.7%; Score 59.6; DB 3; Length 6256;
Best Local Similarity 51.9%; Pred. No. 4.2e-08;
Matches 134; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 313 tatgtgaacctctcaaaaacgaagattcaaggaactatactcttggtgcttgaagaatctg 372
DB 2806 TTCTTGACCTTGATTGAAATGAAGATCAACAGAAAGCATTCCAAAGATATTGGCAATCTT 2865
QY 373 aagaagctcaatcaagtttgatctgtacaacacaatctcacccgggaaatcccatcttct 432
DB 2866 ATTGGCTTACAACTCTCTATCTCTGCAACAAATTTCAGAGGCTACTTCATCTCATCTG 2925
QY 433 ttgggaaatgaagctactgttttttgcgcttaagcgaagaacgcgattgacggtctct 492
DB 2926 TTGGGACAGCTTAGAAACTTAGGCATTTCTAGTCCTACGAAACAACTTAGAGGGGTTTCG 2985
QY 493 attccatagaaactcaagcttattcaagcttaagttgttgatgtctcaaggaatgat 552
DB 2986 ATCCCATTTGGCCATGAGAAATCTTACGAACTTAATATCTTACTGCTGGCACCACAA 3045
QY 553 ttgtgtggaacaattcca 570
DB 3046 TTCAGTGGTTGGATACCA 3063

RESULT 6
US-08-567-375-1
Sequence 1, Application US/08567375
Patent No. 5952485
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.

APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/567,375
FILING DATE: 04-DEC-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645
FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058930
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1648..4383, 5178..5513)
OTHER INFORMATION: /product= "RRK-F"
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
OTHER INFORMATION: resistance gene RRK-F from rice
OTHER INFORMATION: (Oryza sativa)"
US-08-567-375-1

Query Match 6.7%; Score 59.6; DB 4; Length 6256;
Best Local Similarity 51.9%; Pred. No. 4.2e-08;
Matches 134; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 313 tatgtgaacctctcaaaaacgaagattcaaggaactatactcttggtgcttgaagaatctg 372
DB 2806 TTCTTGACCTTGATTGAAATGAAGATCAACAGAAAGCATTCCAAAGATATTGGCAATCTT 2865
QY 373 aagaagctcaatcaagtttgatctgtacaacacaatctcacccgggaaatcccatcttct 432
DB 2866 ATTGGCTTACAACTCTCTATCTCTGCAACAAATTTCAGAGGCTACTTCATCTCATCTG 2925
QY 433 ttgggaaatgaagctactgttttttgcgcttaagcgaagaacgcgattgacggtctct 492
DB 2926 TTGGGACAGCTTAGAAACTTAGGCATTTCTAGTCCTACGAAACAACTTAGAGGGGTTTCG 2985
QY 493 attccatagaaactcaagcttattcaagcttaagttgttgatgtctcaaggaatgat 552
DB 553 ttgtgtggaacaattcca 570
DB 3046 TTCAGTGGTTGGATACCA 3063

Db	2986	ATCCCATGGGCCATAGGAATACTTACTGAACCTTAATATCTTACTGCTGGCACCACAAA	3045
QY	553	ttgtgtggaacaattcca	570
Db	3046	TTTCAGTGGTTGGATACCA	3063

RESULT 7

```

Sequence 1, Application US/08587/680A
Patent No. 5977434
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,680A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,891
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,645
FILING DATE: 29-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/567,375
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-058940US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1648..4383, 5178..5513)
OTHER INFORMATION: /product="RRK-F"
OTHER INFORMATION: /note="xa21 Xanthomonas spp. disease
resistance gene RRK-F from rice (Oryza
sativa)"
US-08-587-680A-1

```

Query Match	6.7%;	Score 59.6;	DB 4;	Length 6256;
Best Local Similarity	51.9%;	Pred. No. 4.2e-08;		

Matches 134; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Accession	Sequence	Position
OY	tatgtgaaacttcacaaaaacggagttcaaggaacataccctctcgagcttggaaatctg	372
Db	TTTCTTGCACCTTGCTATTGTAATTAAGATCACAAGCAGCATTTCCAAAGAGATTGGCAGATT	2865
OY	aagsgcttaactcagtttggatctgtcaacaacaactcaccgggaaatcccatctct	432
Db	ATTGGCTTACACATCTCTTCTCTCAACACAAATTTCAGAGGCTCATCTTCATCATCG	2925
OY	ttgggaaacttgaagtcactgttttttggcgcttaagaaacagattgacggtctct	492
Db	TTGGGCAAGCTTAGAAACTTAGGCATTCTTAGTGCCTACGAAAAACAATTGAGCGGTTG	2985
OY	attccctagaagaaactcacagttattcaagcccthaaagtgttgatgctcaggyaagt	552
Db	ATCCCATTTGGCCATAGGAATCTTACTGAACCTAATATCTTACTGCTGGCACCAACAA	3045
OY	ttgtgtggaacaattcca	570
Db	TTCACTGGTGGATTACCA	3063

RESULT

```

US-08-567-375-15
: Sequence 15, Application US/08567375
: Patent No. 5952485
:
: GENERAL INFORMATION:
: APPLICANT: Ronald, Pamela C.
: APPLICANT: Wang, Guo-Liang
: APPLICANT: Song, Wen-Yang
: APPLICANT: Szabo, Veronique
: TITLE OF INVENTION: Procedures and Materials for Confering
: TITLE OF INVENTION: Disease Resistance in Plants
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/567,375
: FILING DATE: 04-DEC-1995
: CLASSIFICATION: 800
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/004,645
: FILING DATE: 29-SEP-1995
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/475,891
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/373,375
: FILING DATE: 17-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34,774
: REFERENCE/DOCKET NUMBER: 023070-058930
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
:
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 831 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```



```

;
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4091
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 898..3489
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 967..3486
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 898..966
;
US-08-666-271-1

```

```

Query Match
Best Local Similarity 6.0%; Score 53.8; DB 4; Length 3905;
Matches 136; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

```

QY 315 tagtgaactctcaaaaacagatccaaggaactatcctcttgagcttgaaatcga 374
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2913 TATCATCTCTCAAGAAACAGATTGGAAGTCATATTCACGATTTATGAGATCTGT 2972
QY 375 gagctcaatcagttggatctgtacaacaacatctcaccgggaatcccatcttt 434
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2973 TGGACTCTGACGTTGAATCTGTCTCACAACTCTTGAAGGTCATATACCGGATCAT 3032
QY 435 gggaaatgaagtcactgttttttgcggttaacgaaacccgattgaccgtctat 494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3033 TCATAATTATATCATGATCTGATCTTGGATCTCTCATCTAATAAATCAGCGAATA 3092
QY 495 tcttagaagactcaacagttattcaagccttaagtgttgatgtctgaaggaaatgtt 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3093 TCCGACGAGCTTGATCCCTCATCTCTTGAAGTCTTAATCTCTCACAAATCAATCT 3152
QY 555 gtgtgaacattccagtagaagaccctttga 587
    ||||| ||||| ||||| ||||| |||||
DB 3153 TGTGATGATCATCCCAAGGAATAATTTGA 3185

```

```

RESULT 11
US-08-473-553A-1
; Sequence 1, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavate Nuclear Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hombach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELETYPE: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
;
US-08-473-553A-1

```

```

Query Match
Best Local Similarity 5.9%; Score 52.4; DB 3; Length 5733;
Matches 125; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

```

```

QY 324 ctcaaaaacgagatccaaggaactatcctctgagcttggaatctgaaatcga 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3120 CTACTCAACAGCTACACCGGGGTGTTCCAGCGAGTGGTGTAAACAAGCTTGA 3179
QY 384 cagttgagatctgtacaacaacatctcaccgggaatcccatctcttggaat 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3180 GATCCGACATGCGGAGCTGACCTGACACCGGAGAGATTCCGACGATTAAGTAAC 3239
QY 444 gaagtcactgttttttgcggttaacgaaacccgattgaccgtctatctcctaga 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3240 GAACATCTCACTCTGTTCTTCTTACATCAACACTTAACCGGTATATACCGGA 3299
QY 504 actcaagttattcaagccttaagtgttgatgtctcaggaagattgtgtgaac 563
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3300 GCTTCCGCTTATGACCTGGAATCTCTGATTTATCAATCAATGATTAACCGAGA 3359
QY 564 aattcc 569
    |||||
DB 3360 AATCCC 3365

```

```

RESULT 12
US-08-238-163-3
; Sequence 3, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAVITCH, John M.
; APPLICANT: POWELL, Ann
; APPLICANT: STOTZ, Henrik
; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
; TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,163
FILING DATE: 03-MAY-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 2307E-540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 421..1401
US-08-238-163-3

Query Match 4.5%; Score 40.6; DB 1; Length 2075;
Best Local Similarity 48.5%; Pred. No. 0.01;
Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 343 ggaactatccttcgagcttggaatctgaagagctcaatcagtttgatctgtaaac 402
DB 733 ggaacattccactgacgttcgagctcacaaatctcaaatctgaagctcagcttc 792
QY 403 aacaaatccacgggaaatccatctcttgggaaatgaagctcagtttttttg 462
DB 793 actaaccttaccagctccatccctgaattccttaccagctgaagaaatttaccgttc 852
QY 463 cggcttaacgaaacgagttgacgggtccatctcctctagaagaactcagttattcaagc 522
DB 853 gattgaattacacatatttaccggaacatcccttccctctccagcttccggaat 912
QY 523 cctaaagttgtagtgcacgggaatgattgtgtggaacaattcagta 573
DB 913 ttccraccatgacttaccatctgaatcgaacacacacggaacatccgga 963

RESULT 13
US-08-244-646-14
Sequence 14, Application US/08244646
Patent No. 5744692
GENERAL INFORMATION:
APPLICANT: Cervone, Felice
APPLICANT: De Lorenzo, Giulia
APPLICANT: Salvi, Giovanni
APPLICANT: Albersheim, Peter
APPLICANT: Darvill, Alan
APPLICANT: Bergmann, Carl
TITLE OF INVENTION: Nucleotide Sequences Coding An
TITLE OF INVENTION: Endopolygalacturonase Inhibitor
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sally A. Sullivan
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,646
FILING DATE: 06-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM 91A 000915
FILING DATE: 06-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/IT/00158
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 19-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1116 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Phaseolus vulgaris
STRAIN: Saxa
IMMEDIATE SOURCE:
CLONE: lambda pGIP-3.3
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1026
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1027..1116
US-08-244-646-14

Query Match 4.0%; Score 36; DB 2; Length 1116;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 90; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 319 gaactcaaaaagagattcaaggaactatacttctgagcttggaatctgaaggt 378
DB 835 gatctgaggaaacgctgattggaacgctcagggacgtaacgacgttaagttt 894
QY 379 ctataagttgagctgtacacacaaactcaccgggaaatccatcttcttgga 438
DB 895 ctgcagaatttaatttgacttcaacatctgtcgggagatcttccacaggtggaaac 954
QY 439 aaattgaagctactgttttttgcggttaacgaaacgagttaccggtccattcct 498
DB 955 ttgaaaggttgacgtttcttcttaccacaacaaagtgctgtggttcttctctt 1014

RESULT 14
US-08-968-563-5
Sequence 5, Application US/08968563
Patent No. 6013494
GENERAL INFORMATION:
APPLICANT: CHARLES E. NAKAMURA
APPLICANT: ANTHONY A. GATEWAY
APPLICANT: AMY (KUNANG-HUA) HSU
APPLICANT: RICHARD D. LA REAU
APPLICANT: SHARON L. HAYNIE
APPLICANT: MARIA DIAZ-TORRES
APPLICANT: DONALD E. TRIMBUR
APPLICANT: GREGORY M. WHITED
APPLICANT: VASANTHA NAGARAJAN

```

; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 4 CAMBRIDGE PLACE
; STREET: 1870 SOUTH WINTON ROAD
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,563
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: NOVEMBER 13, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9982
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1380 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: GPD1
; US-08-968-563-5

```

```

Query Match          4.0%; Score 35.6; DB 5; Length 1380;
Best Local Similarity 58.5%; Pred. No. 0.25;
Matches 62; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 449 cactgttttttcgagcctaagaacacgattgacggcttcctattcctagaagactca 508
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 479 CACATCATTTTTCGCCCGTATCTGTAGCCATGGAAGGTGATGATTCACACGTCA 538
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 509 cagtattccaagccttaagttgtgatgtctcaagggaatgattt 554
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 539 GAGCTATCTCCTGTCTAAGGGTTTGAGTTGAGTGTGCTTAAGGTCGT 584

```

```

RESULT 15
5231168-1/c
; Patent No. 5231168
; APPLICANT: DIEGIEL, MORTEN,BORRE, MARTIN,JEPSEN, SOREN,
; VUUST, JENS,RIENCK, KLAUS,WIND, ANNETTE,JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989

```

```

; SEQ ID NO:1:
; LENGTH: 3095
5231168-1

```

```

Query Match          4.0%; Score 35.4; DB 7; Length 3095;
Best Local Similarity 53.2%; Pred. No. 0.43;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 741 tgtttgtaactcataagaacctgttcaatgatatataagaagcttgagagacagt 800
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2495 TATTTCTATATATACAGAAATCCTAGTATCTTTATATATATATATATATATAT 2436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 801 aacttgcaatgataatgtatgtatgtatgagaagaagtgaatgagaaatgttgtaattga 860
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2435 ATAAATATTAGTTCATATATATAGAAAAATTTATTTATATATTTTATTTT 2376
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 861 ttgtgttctctatgtaactt 881
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2375 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 2355

```

```

Search completed: June 23, 2000, 22:46:00
Job time: 40306 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:38 ; Search time 6198.48 Seconds
(without alignments)
584.592 Million cell updates/sec

Title: US-09-180-798-28
Perfect score: 894
Sequence: 1 ggaacgcatcaagctccga.....gttaacttgattcttattta 894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	419.8	47.0	450	20	T04109	T04109 59 Lambda-P
2	387.4	43.3	443	63	A1999615	A1999615 70156713
3	265	29.6	556	23	R89908	R89908 16353 Lambd
4	256.4	28.7	353	23	H37300	H37300 15429 Lambd
5	223.8	25.0	534	33	AA394359	AA394359 25942 Lam
6	220.6	24.7	608	81	AW443205	AW443205 EST308135
7	219.6	24.6	720	64	AM031110	AM031110 EST274417
8	213.8	23.9	430	79	AM307218	AM307218 sf54c07.Y
9	210.4	23.5	657	80	AM350720	AM350720 GM210009A
10	209.6	23.4	570	50	A1676939	A1676939 605047a07
11	208.6	23.3	541	64	AM036855	AM036855 614019G01
12	207.8	23.2	637	80	AM350549	AM350549 GM210009A
13	205	22.9	646	74	AM221939	AM221939 EST298750
14	204.2	22.8	545	64	AM037836	AM037836 EST279465
15	202.6	22.7	505	25	N65416	N65416 20456 Lambd
16	199.2	22.3	511	74	AM221278	AM221278 EST297747
17	197.6	22.1	714	64	AM030188	AM030188 EST273443
18	196.8	22.0	364	42	A1100683	A1100683 33804 Lam
19	194.8	21.8	579	43	AU031292	AU031292 AU031292
20	189.4	21.2	417	21	D46231	D46231 R1C510770A
21	186.6	20.9	616	42	AF074734	AF074734 AF074734
22	185.2	20.7	348	42	A1100682	A1100682 33803 Lam
23	184	20.6	695	64	AM038168	AM038168 EST279825
24	183.4	20.5	460	36	C72525	C72525 C72525 R1C6
25	182.2	20.4	562	50	A1691407	A1691407 606015E09
26	175.4	19.6	599	74	AM219797	AM219797 EST302279
27	173.2	19.4	334	23	H37195	H37195 15324 Lambd
28	171.8	19.2	599	63	AM011134	AM011134 SP17B03 P
29	171	19.1	339	42	A1100679	A1100679 33800 Lam
30	168.2	18.8	469	63	A1992868	A1992868 701493826
31	167	18.7	676	74	AM220075	AM220075 EST302558
32	163	18.2	569	59	A1775448	A1775448 EST265348
33	156.6	17.5	423	35	C22371	C22371 C22371 R1C6
34	155.6	17.4	336	23	H37296	H37296 15425 Lambd
35	154	17.2	401	37	AA712221	AA712221 31949 Lam
36	151.6	17.0	612	47	A1496325	A1496325 SB05C09.Y
37	149.4	16.7	458	59	A1779982	A1779982 EST258861
38	149.4	16.7	463	74	AM221200	AM221200 EST297669
39	149.2	16.7	502	79	AM286582	AM286582 L61_334_LF
40	148.8	16.6	514	35	C22372	C22372 C22372 R1C6
41	147.8	16.5	368	64	AM032898	AM032898 EST276457
42	147.8	16.5	423	59	A1774881	A1774881 EST255981
43	147.8	16.5	423	62	A1896737	A1896737 EST266180
44	147.8	16.5	488	64	AM031917	AM031917 EST275371
45	147.8	16.5	497	59	A1772079	A1772079 EST253179

ALIGNMENTS

RESULT	1	LOCUS	450 bp	EST	06-NOV-1997
DEFINITION	T04109	59 Lambda-PRL1	Arabidopsis thaliana	CDNA clone	SCH97P, mRNA
ACCESSION	T04109	GI:315269			
VERSION	EST				
KEYWORDS	EST				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1	(bases 1 to 450)			
AUTHORS	Newman,F., deBrujn,F.J., Green,P., Keegstra,K., Kende,H., McIltoosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.				
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones				
JOURNAL	Plant Physiol.	106,	1241-1255	(1994)	

MEDLINE	95148729.
COMMENT	Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing MI Tel: 517-353-0854 Fax: 517-353-9168 Email: 22313cne@bm.ci.msu.edu.
FEATURES	Location/Qualifiers
source	1..450
	/organism="Arabidopsis thaliana"
	/strain="var columbia"
	/db_xref="taxon:3702"
	/clone="SCH97P"
	/note="Vector: Lambda Shlox-1; Site_1: EcoRI; Site_2: HindIII; Lambda PRL1 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The library was made in Novagen's Lambda Shlox-1 with (oligo dt primed) directional inserts cloned between the EcoRI and HindIII sites."
BASE COUNT	125 a 94 c 81 g 147 t 3 others
ORIGIN	
Query Match	47.0%; Score 419.8; DB 20; Length 450;
Best Local Similarity	98.4%; Pred. No. 3.6e-92;
Matches	443; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
QY	24 ggaagctcttcacgcgcttcgcgcgagcttacagatccagaacatgtttcagagctg 83
DB	1 GCAGCTCTTCACGCCCTTCGCCGAGCTTATCAATCAATGTTGTTCAGAGTGG 60
QY	84 ggaatcaactctgttaactcctgtactgtgttcatgtcactgttaatacaacacatca 143
DB	61 GGATCCACTCTTGTAACTCTTGACTGTGTTTATGTCACCTTTATCAACACATCA 120
QY	144 agtcaactcgtctggaatttgggaattcaactatctggaacttagtacctggaacttg 203
DB	121 AGTCACTCGCTGATTTGGGGAATCAACTTACTGTGATCTGATCACTGAACTGG 180
QY	204 gaagcttgaacatttacaatctatctgtagaatactaatcttggccttttgatattct 263
DB	181 GAACTTGAACATTATCAATCTGTATGAAATCAATCACTTTTGCTTTGATTATCT 240
QY	264 gaaacatttaccattacagtcacacataataacatttgccttgccttgccttgccttgc 323
DB	241 GAAACATTACATTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY	324 ctacaaaacagagatcaaggaactatccttctgagcttggaaatctgaagatctaat 383
DB	301 CTACAAAACAGATCTCAAGGAACTATACCTTCTGAGCTTGAATCTGAAATCTGAA 360
QY	384 cagtttgatctgtacaaacacatctcac--gggaatctccatcttctt--gggaataa 441
DB	361 CAGTTTGATCTGTACAAACACATCTCACCGGGAATAATCCATCTTCTTGGGAAAA 420
QY	442 ttgaagcactgttttttgcgactaac 471
DB	421 TTGAAGNCACTGTTTGTTCGCGNTTTC 450
RESULT	2
LOCUS	A1999615
DEFINITION	701556713 A. thaliana, Columbia Col-0, rosette-3 Arabidopsis thaliana cDNA clone 701556713, mRNA sequence.
ACCESSION	A1999615

VERSION A1999615.1 GI:5846520
 EST.
 KEYWORDS thale cress.
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 443)
 AUTHORS Chen, J., Montiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Broska, P., Gorgone, G., Burns, D., Griffin, J., Monanoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, I., Policky, J., Suzuki, G., Argente, C., Shah, S., Nobriza, A., Murry, L., Turner, C., Kikorian, S., Elder, L. and Hanson, D.

TITLE Arabidopsis thaliana Gene Expression Microarray
 JOURNAL Unpublished (1999)
 COMMENT On Mar 10, 1998 this sequence version replaced gi:2948106.
 Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES
 SOURCE
 1..443
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701556713"
 /clone_1lb="A. thaliana, Columbia Col-0, rosette-3"
 /tissue_type="rosette"
 /dev_stage="4 - 7 weeks"
 /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

BASE COUNT 135 a 88 c 54 g 124 t 42 others
 ORIGIN

Query Match 43.3%; Score 387.4; DB 63; Length 443;
 Best Local Similarity 90.2%; Pred. No. 2.6e-84;
 Matches 388; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 455 ttttttcggttaacgaacacgattgacggctctattctcagagaaactcacagttta 514
 |||||
 DB 430 TTTTTCGGGCTNACGAAAACNATNGACNGTNTATTCCTAGAGNACTCAGATT 371
 |||||
 OY 515 ttccaagccttaagttgtgtcctcaggaatgattgtgtggaacaactccagtag 574
 |||||
 DB 370 TTTCAAGCCNTAAANTNGTGTATGATCTCAGAGATGATNGTGTGNAACAATCCAGTAG 311
 |||||
 OY 575 aaggacctttgaacacactcctatgacaaactttgagaacaaactgagattgaggagac 634
 |||||
 DB 310 AAGGACNTTTTGNANACATTCCTAGCNAACCTTGAACNCAACCTAGATGAGAGGAC 251
 |||||
 OY 635 cagaactactaggtcttgcagctatgacacaaattgcaacttaaaagaagtgaagac 694
 |||||
 DB 250 CGAAGCTACTAGCTTCGAGNTNTGACACCAATTCGCACTTNAAGAAGTTGAAGAAC 191
 |||||
 OY 695 ctataaagaagaattgtgagacttgaagaactgtgacaaagtgttgtaactta 754
 |||||
 DB 190 CTNTAANGAAGATGTAGGTGACNTTGTAAGAANTCTGTACCAAGNTTTTGTATTCCTA 131
 |||||

OY 755 tataagccttgttcatctatatatgaagccttggagacagtaactgcacatgtat 814
 |||||
 DB 130 TATAGACNCTGTTTCATGTATATGAAACCTTGAGAGCACTAANTTGCACATGTAT 71
 |||||
 OY 815 tggatgtggaagaacatgaaatgagattgcttttaattgattgtttcttat 874
 |||||
 DB 70 TNNTANTNGTAGNANAGATTGNATGACATTCCTTTGTAATGATNTGTCTTTCTTAT 11
 |||||

OY 875 gtaacttgaa 884
 |||||
 DB 10 GTAANTGAA 1
 |||||

RESULT 3
 R89998 556 bp mRNA EST 30-DEC-1997
 LOCUS 16353 Lambda-PRL2 Arabidopsis thaliana cDNA clone 187H577, mRNA
 DEFINITION
 ACCESSION R89998
 VERSION R89998.1 GI:957538
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 556)
 AUTHORS Newman, J., deBruin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Onltroge, J., Raiknel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.
 TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 JOURNAL plant physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729

COMMENT On Apr 14, 1993 this sequence version replaced gi:693023.
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tch@dm.cl.msu.edu
 Seq primer: T7 dye primer.

FEATURES
 SOURCE
 1..556
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="187H577"
 /clone_1lb="lambda-PRL2"
 /note="Vector: lambda Z1p-lox; Site_1: SalI; Site_2: NotI; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark - rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Z1p-lox. The cDNA inserts were directionally cloned with SalI-Not arms using oligo dT primed cDNA."

BASE COUNT 151 a 118 c 123 g 146 t 18 others
 ORIGIN

Query Match 29.6%; Score 265; DB 23; Length 556;
 Best Local Similarity 80.0%; Pred. No. 1.6e-54;
 Matches 335; Conservative 0; Mismatches 79; Indels 5; Gaps 2;

OY 313 tatagtgaactctcaaaacgagattccaaggaactatactcttgaagttggaatctg 372
 |||||
 DB 49 TATCTAGAGCTCTCAAAAACACATCCAGAGACTATACCTTCGAACTTGGAATCTCG 108
 |||||

OY	493	atctcctagagaaactcacagttattctcaagccctaaagttgttgatgcttcagggagatcat	552
Db	354	ATTCCCAAGGAACGTCTGCTGTTTCAAGCCTTTAAAGTAGTGAGTCTCCAAACATGAT	295
OY	553	ttgtgtggaacattccacagtagaagacctttgacaacattccctatgcaaaacttgag	612
Db	294	TTATGTGGTACATCTCCACACTCTGGGCGATTGCGACATATTCCATTGAAATTAACCTTGAG	235
OY	613	aacaacactgagattggaaggagacgaactactaggtcttcgagctatgacacccaattgc	672
Db	234	AATATATCCCCGCGTTGGAAAGTCCACAGATTGTTGGACACTAGTAAGTTATGACACCAAACTGC	175
OY	673	acttaaaa	681
Db	174	TCGTGAATTA	166

RESULT	10			
LOCUS	A1676939/c			
DEFINITION	A1676939	570 bp	mRNA	EST
ACCESSION	605004/AF07.x1	605	- EndospERM	CDNA library from Schmidt lab
VERSION	A1676939			Zea mays
KEYWORDS	A1676939.1	GI:4885819		
SOURCE	EST.			
ORGANISM	Zea mays.			
	Zea mays			

REFERENCE	1 (bases 1 to 570)
AUTHORS	Walbot,V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	University
COMMENT	Unpublished (1999) On Jun 5, 1998 this sequence version replaced gi:3189462.

JOURNAL Unpublished (1999)
COMMENT On Jun 5, 1998 this sequence version replaced g1:3189462.
Contact: Walpot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walpot@stanford.edu
Plate: 605047 row: A column: 07.
Location/Qualifiers
FEATURES
source 1..570

FEATURES	SOURCE	Location/Qualifiers
	1.	.570
		/organism="Zea mays"
		/cultivar="Ohio43"
		/db_xref="taxon:4577"
		/clone_lib="605 - Endosperm cDNA library from Schmidt lab"
		/tissue_type="nucellar, embryo, and endosperm"
		/dev_stage="10-14 days post pollination"
		/lab_host="DH5(alpha)"
		/note="Organ: kernel; Vector: PAD-GA4-2; Site:1: EcorI; Site:2: XhoI; kernel endosperm cDNA library from Schmidt lab"
BASE COUNT	150 a	131 c 127 g 162 t
ORIGIN		

Query Match	23.4%	Score 209.6;	DB 50;	Length 570;
Best Local Similarity	73.9%	Pred. No. 4.8e-41;		
Matches 266; Conservative	0;	Mismatches 94;	Indels 0;	Gaps 0;

QY	313	tataatgaactacacaanaacgaagatctcaaggaactatacccttcgagcttggaatctg	372
Db	541	TATCTGGAACCTTAAATAGATACACTTCAGGAACAATCCCTTCGAGGTTGGCAACTTG	482
QY	373	aagagctctaatcagtttgagatctgtaacaacaacatctccaggggaaaaatcccatctct	432
Db	481	AAGATCTAAATAGGTTCCACTTGTACAGAACCAACATTCGGGGACATATACCTCTGGC	422

Oy 433 ttgggaataatgaatcaactgttttttgggggttaagcaaaacgattgacggfct 492
 Db 421 CTGGAAAGTtGAAgtCCCTtGtATtCTtGGGtCTCAATGCAATCAATTtGACTGGGCC 362
 Oy 493 attccttagagacatcacagtatttccaagccctaagttgtgtgtctcaaggaatgat 552
 Db 361 ATCCCAAGCGAACTCGGTGGAATATCTAGTCTAAATGTGTGATGTTCACGAATGAT 302
 Oy 553 ttgtgttgacaattccagtagaaggaaccttggacaacatccctatgcaaaacttgag 612
 Db 301 TTGTGCGGGGAATATCCACGCTCTGGACATTTAGCATATTTCTTCAGACCACTTGAG 242
 Oy 613 aacaacctgagattggagggagcagaactactagttgttcggagctatgacaacaattgc 672
 Db 241 AAGAACCCGCGCTTGGAAAGTCTCAGAGCTTACAAAGCGCTGCTCATATATCGACCACTGC 182

RESULT	11
AM036865	
LOCUS	AM036865 541 bp mRNA EST 15-SEP-1999
DEFINITION	G1401G10.Y2 G14 - root CDNA library from Walpole Island Zea mays CDNA, mRNA sequence.
ACCESSION	AM036865
VERSION	AM036865.1 GI:5895619
KEYWORDS	EST.
SOURCE	Zea mays.
ORGANISM	Zea mays.

REFERENCE	1 (bases 1 to 541)
AUTHORS	Walbot,V.
TITLE	Matze ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL	Unpublished (1999)
COMMENT	On Jul 7, 1999 this sequence version replaced gi:5407058.

JOURNAL
COMMENT
Unpublished (1999)
On Jul 7, 1999 this sequence version replaced g1.5407058.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614019 row: G column: 10.
Location/Qualifiers
1. .541

FEATURES	Source
Place: 014013 106: 9	Location/Qualifiers
	1. .541
	/organism="Zea mays"
	/cultivar="W23"
	/db_xref="taxon:4577"
	/clone_id="614 - root cDNA library from Walbot Lab"
	/tissue_type="root"
	/dev_stage="3-4 days old"
	/lab_host="XLDR"
	/note="Organ: root; Vector: pBlueScriptII SK+; Site: 1; Ecoli: Site_2; XhoI: 3-4 days old root tissue from Walbot Lab (LW)"
BASE COUNT	154 a 121 c 126 g 139 t 1 others
ORIGIN	

Query Match	23.3%	Score 208.6;	DB 64;	Length 541;
Best Local Similarity	73.6%	Pred. No. 8.3e-41;		
Matches 265; Conservative	0;	Mismatches 95;	Indels 0;	Gaps 0;

OY	313	tatagttgacctctcaaaaaacgagattcgaagaacttactcttgtagcttggaacctg	372
Dd	35	TATCGGACCTTATAAACAATACATTCAAGAACAACTCCCTTGACTTCGGCAATTG	94
OY	373	aagagtctaatacgtttggatctgcacaacaatctcacgaggaaatcccatcttc	432
Dd	95	AAGAATCTAATTAAGCTTGACTGTGCACGAACAACATTTGGGGACCAATACCTCCTGCC	154
OY	433	ttgggaaaattgaagtcacttgttttttttggcggtttaacgaanaacgatitgaccggtct	492

```

Db 155 CTTGAAGAGTTGAAGTCCCTTGTATCTTCTGCTCAAGGCAATATTGACTGGGCC 214
QY 493 attctagaagaccacagattattccaagcctaaagtgtgatgtctcaaggaaatgat 552
Db 215 ATCCCAAGGGAACACGCTGGAAATCTAAGTCTCAAACTGNTGATGTTTCAGTAATGAT 274
QY 553 ttgtgtgaacaactccaagtagaagacctttgaaacacattcctatgcaaaactttgag 612
Db 275 TTGTGCGGGAGAGATTCCTCAGCTCTGACCATTTGAGCATTTCTCTGAGCACTTTGAG 334
QY 613 aacacacgagattggagagacacgaactactagcttctgcgagctcatcaaccacattgc 672
Db 335 AAGAAACCGCCTTGGAAGGCTCAGACCTACAGGCTGCGCATATACGACCACTTCGC 394

```

RESULT 12

AM350549

```

LOCUS GM210009A10A12R Gm-r1021 Glycine max cDNA clone Gm-r1021-3095 3',
DEFINITION

```

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers

FEATURES
source

1..646
/organism="Lycopersicon esculentum"
/cultivar="T496"
/db_xref="taxon:4081"
/clone_1db="tomato fruit red ripe, TAMU"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 181 a 126 c 138 g 201 t
ORIGIN

Query Match 22.9%; Score 205; DB 74; Length 646;
Best Local Similarity 64.6%; Pred. No. 6.3e-40;
Matches 374; Conservative 0; Mismatches 115; Indels 90; Gaps 1;

QY 16 tccgaagggaagcgtcttcacgcgcttcgcggaagatcatagatccagacaatgtgtt 75
DB 158 TCAGAGGGAGGAGCTTTGTAGCCCTCCGCGAGCTTATCTGACCCGGTAAAGCTGTA 217
QY 76 cagagattggatccacacctgttaattctgtactgtgtctgtatgtcatgtatca 135
DB 218 CAGAGCTGGAGTCCAAATCTGTATACCTGTACTGTTCATGTCATCTGCACAGGA 277
QY 136 caccatcaagtcacgtcgtctgatttgggaattcaactatctgacatctagta 195
DB 278 GATATATCAAGTACTCGTGATGATCTTGGAACTCAAGTATATCGTCAATTGGTACT 337
QY 196 gaacttgggaagcttgacaatataatgtatgtatgtatgaatcaatcatttgc 255
DB 338 GAGCTCGGAAACCTTGACATCTACAGTATCTG----- 371
QY 256 gattatctgaacaattacattacatcagtcacacatatataatttggatcatat 315
DB 371 ----- 371
QY 316 agtgaactcacaaaaacgagatccaaggaactatcctctgagcttggaaatcaga 375
DB 371 ---GAGCTTTACAAAATAATATTCAGGAAACCATCCCTAAGAGCTCGTAACTTGAG 427
QY 376 agtcaatcagtttgatctgtacacaacaatctcacgcggaatccatcttcttg 435
DB 428 AGCTTATTAAGTCTGATCTGTACACACAAATATTTGGGACAAATTCCTACTTCACT 487
QY 436 ggaagaattgaagtcactgttttttcgagcttaacgaacaacgagttgacggtct 495
DB 488 GGAAGAACTGAAAACCTTGTCTTTCGCTTAATGATTAACAACCTAACAAGACAAATC 547
QY 496 cctagagaactcacagttatccaagccttaagtgatgtatgtatcagggaaatgt 555
DB 548 CCAAGAGAACTTACAGCTTTCTAGCCGAAAGTTGGAGATGCTCAATTAATGATTGG 607
QY 556 tctgagaactccagtagaaggaacttggacaatt 594
DB 608 TGTGGAACAATTCCTACTTCTGTCTCATTTTGACATATT 646

RESULT 14
LOCUS AM037836 545 bp mRNA EST 18-OCT-1999
DEFINITION EST279465 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

ACCESSION AM037836
VERSION AM037836.1 GI:5896590
KEYWORDS EST.
SOURCE
ORGANISM

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 545)
D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J., Roming, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Niernann, W., Fraser, C. W., Venter, J. C., Martin, G. B., Tanksley, S. D. and Giovannoni, J.
Generation of ESTs from tomato leaf tissue
Unpublished (1999)
On Jun 22, 1998 this sequence version replaced gi:3246851.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU

TITLE
JOURNAL
COMMENT

FEATURES
source

1..545
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande Flor"
/db_xref="taxon:4081"
/clone="cLEF3j21"
/clone_1db="tomato mixed elicitor, BTI"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; CLEF - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisocoumaric acid, BPH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

BASE COUNT 159 a 118 c 110 g 158 t
ORIGIN

Query Match 22.8%; Score 204.2; DB 64; Length 545;
Best Local Similarity 70.0%; Pred. NO. 9.8e-40;
Matches 275; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 313 tataggaactctacaaaaacgagatccaaggaactatcctctgagcttggaaatcag 372
DB 68 TATCTGAGCTTTACAAAATAATATTCAGGAAACCATCCCTAAGAGCTCGTAACTTGAG 127
QY 373 aagagtcataatcagtttgatctgtacacaacaatctcacgcggaatccatcttct 432
DB 128 AAGAGCTTATTAAGTCTGTGATCTGTACACAAATATTTGGGACAAATTCCTACTTCA 187
QY 433 ttgggaattgaagtcactgttttttcgagcttaacgaacaacgagttgacggtct 492
DB 188 CTGGAAGAACTGAAAACCTTGTCTTTCGCTTAATGATTAACAACCTAACAAGACAAATC 247
QY 493 attcctagaactcacagttatccaagccttaagtgatgtatgtatcagggaaatgt 552
DB 248 ATCCCAAGAGAACTTCTAGCATTTCTACCCGAAAGTTGGAGATGCTCAATTAATGAT 307
QY 553 ttgtggaacattccagtagaaggaacttggacaatttggacaatttggacaatttgg 612
DB 308 TTGTGTGGAACAATTCCTACTTCTGTCTCATTTTGACATATTTCAG 367
QY 613 aacaactgagattggaggacagaactactagttcttgcagctatgacacaacttgc 672
DB 368 CACAATCTCGAGTGAAGTCCGAGTGTGTTGGAGCTCGCTACCTAGACACCAACTGC 427

Qy 673 acttaaaagaagttgaagaacctataagaag 705
Db 428 TCTTGAAGCTTGATGATTAAGCTTCAGAGAG 460

RESULT 15
LOCUS N65416
DEFINITION 20436 Lambda-PRL2 Arabidopsis thaliana cDNA clone 228M97, mRNA sequence.
ACCESSION N65416
VERSION N65416.1 GI:1217042
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS Newman,F., deBrujn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Onltogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On May 8, 1995 this sequence version replaced gi:801103.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@dm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 305
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="228M97"
/note="Vector: lambda Z1p-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda Z1p-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

BASE COUNT 81 a 68 c 63 g 81 t 12 others
ORIGIN

Query Match 22.7%; Score 202.6; DB 25; Length 305;
Best Local Similarity 78.3%; Pred. No. 2.3e-39;
Matches 235; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 382 atcagttgagctctgtacaacaacatctcacgcgggaaatcccatctcttttgggaaaa 441
Db 1 ATCAGCTTGCTGTGTCAACAACAATCTACAGGGATAGTCCCACTTWTTCGGAAAA 60

Qy 442 ttgaagtcacgtgttttttggcggttaacgaacacgattgacgcgtctctattctctaga 501
Db 61 TTNAAGTTCGTGCTTTTAACGCGCTTAATGACACCGATTGANGGGCCAAATCCCTAGA 120

Qy 502 gaactcacagttattcaagccttaaggttgatgtctcagggaatgatttgttggga 561
Db 241 AGGTTGAGGGGCCCGAATTAATCTGAGTTCGCAAGNTTCGCACTTACTGACCTGAAAA 300

Search completed: June 23, 2000, 19:06:42
Job time: 27435 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:23:23 ; Search time 12463.9 Seconds

(Without alignments)
-82.966 Million cell updates/sec

Title: US-09-180-798-30

Perfect score: 1063
Sequence: 1 tcgaccaccgcgtccgacga.....tcgcaaaaaaaaaaaaaa 1063

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_bal:*
2: gb_bal:*
3: gb_bal:*
4: gb_bal:*
5: gb_bal:*
6: gb_bal:*
7: gb_bal:*
8: gb_bal:*
9: gb_bal:*
10: gb_bal:*
11: gb_bal:*
12: gb_bal:*
13: gb_bal:*
14: gb_bal:*
15: gb_bal:*
16: gb_bal:*
17: gb_bal:*
18: gb_bal:*
19: gb_bal:*
20: gb_bal:*
21: gb_bal:*
22: gb_bal:*
23: gb_bal:*
24: gb_bal:*
25: gb_bal:*
26: gb_bal:*
27: gb_bal:*
28: gb_bal:*
29: gb_bal:*
30: gb_bal:*
31: gb_bal:*
32: gb_bal:*
33: gb_bal:*
34: gb_bal:*
35: gb_bal:*
36: gb_bal:*
37: gb_bal:*
38: gb_bal:*
39: gb_bal:*
40: gb_bal:*
41: gb_bal:*
42: gb_bal:*
43: gb_bal:*
44: gb_bal:*

45: gb_bal:*
46: gb_bal:*
47: gb_bal:*
48: gb_bal:*
49: gb_bal:*
50: gb_bal:*
51: gb_bal:*
52: gb_bal:*
53: gb_bal:*
54: gb_bal:*
55: gb_bal:*
56: gb_bal:*
57: gb_bal:*
58: gb_bal:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1063	100.0	1063	5 A67825	A67825 Sequence 30
2	912.4	85.8	1106	5 A67817	A67817 Sequence 22
3	841.8	79.2	981	5 A67819	A67819 Sequence 24
4	769.6	72.4	789	5 A67821	A67821 Sequence 26
5	361	34.0	894	5 A67823	A67823 Sequence 28
6	211.8	19.9	2089	5 A67827	A67827 Sequence 32
7	155	14.6	936	8 SBU62279	U62279 Sorghum bic
8	103.4	9.7	1755	8 DCU93048	U93048 Daucus caro
9	103.4	9.7	1815	5 A67797	A67797 Sequence 2
10	99	9.3	4604	7 LEIRGEWE	X95269 L. esculentu
11	84.8	8.0	101284	8 ATAC009991	AC009991 Arabidops
12	81.6	7.7	116944	8 AC008017	AC008017 Arabidops
13	81.4	7.7	5033	8 INU77888	U77888 Ipomoea nil
14	79.8	7.5	81513	50 AC021198	AC021198 Arabidops
15	79.2	7.5	134402	8 ATAC011765	AC011765 Arabidops
16	78.8	7.4	100887	50 ATAC011620	AC011620 Arabidops
17	78.8	7.4	108355	50 AC005957	AC005957 Arabidops
18	78.6	7.4	133840	8 AC003981	AC003981 Complete
19	76.8	7.2	113529	44 AC015446	AC015446 Arabidops
20	76.4	7.2	23143	8 F15P11	AF160760 Arabidops
21	75.8	7.1	142418	7 AP000815	AP000815 Oryza sat
22	75.4	7.1	3176	5 E12705	E12705 Arabidops
23	75.4	7.1	3176	8 ATU47029	U47029 Arabidops
24	75	7.1	5940	8 OLUT72727	U72727 Oryza longi
25	74.6	7.0	84440	7 AB010075	AB010075 Arabidops
26	74.6	7.0	118086	7 ATFEH11	AL021684 Arabidops
27	72	6.8	87434	7 AB010698	AB010698 Arabidops
28	71.4	6.7	3573	5 A57133	A57133 Sequence 4
29	71.4	6.7	5368	8 U42445	U42445 Lycopersico
30	71.4	6.7	6471	5 A57130	A57130 Sequence 1
31	71.4	6.7	6471	8 U42444	U42444 Lycopersico
32	71.4	6.7	61712	7 AB011061	AB011061 Arabidops
33	70.8	6.7	4081	5 A67815	A67815 Sequence 20
34	70.8	6.7	6695	5 A67796	A67796 Sequence 1
35	70.8	6.7	98471	8 AC012654	AC012654 Arabidops
36	70.6	6.6	2278	8 SBRUK1	Y14600 Sorghum bic
37	70.6	6.6	2786	7 AB029327	AB029327 Nicotiana
38	70.6	6.6	3979	5 A67428	A67428 Sequence 1
39	70.6	6.6	3979	5 A67429	A67429 Sequence 2
40	70.6	6.6	3979	5 AF053993	AF053993 Lycopersi
41	70.6	6.6	4123	5 A67434	A67434 Sequence 7
42	70.6	6.6	4123	8 AF053998	AF053998 Lycopersi
43	70.2	6.6	9424	8 OSU72724	U72724 Oryza sativ
44	70.2	6.6	52717	7 AB019227	AB019227 Arabidops
45	70	6.6	101284	8 ATAC009991	AC009991 Arabidops

ALIGNMENTS

```

RESULT 1
A67825 1063 bp DNA PAT 05-MAY-1999
LOCUS Sequence 30 from Patent WO9743427.
DEFINITION A67825
ACCESSION A67825
VERSION A67825.1 GI:4756647
KEYWORDS
SOURCE
ORGANISM
unclassified.
unclassified.
REFERENCE 1 (bases 1 to 1063)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICRITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source
location/Qualifiers
1..1063
/organism="unclassified"
/db_xref="taxon:32644"
106..762
/note="unnamed protein product"
/codon_start=-1
/db_xref="GI:4756648"
/translation="MASRNYRWELFASLTLTALIHVLEANSSEDAIYALRSLTDP
DHVLSMDPTLVNPTCFWHTYCNODNRVTRDNLGNSLGHAPLGLKLEHLYLELY
KNNIQGITPSELGNLKNLISLDLYNNNLGIVPTSLGKLSLVFLRLNDRLTGPILR
ALTAIPSLKVVDSNDLCGTIPINGPFAHIPLOFNENNPRLGPELGLASDYDINCT
"
BASE COUNT 313 a 242 c 206 g 302 t
ORIGIN

```

```

Query Match 100.0%; Score 1063; DB 5; Length 1063;
Best Local Similarity 100.0%; Pred. No. 2.3e-263;
Matches 1063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 tcgaccacgcgtccgacgaaccccaatttctccctccatcttcttagaataattact 60
DB 1 TCgACCCACGCGTCGACGAAACCTTAATTTCTTCCTCATCTTGTACAAAATTACT 60
QY 61 caaatlctatagattactctctctcgcaccccgatagatcacatggtctcgaac 120
DB 61 CAATTCCTATTGATTACTCTCTCTCGACCTCCGATGACTACATGGGCTCGAAGC 120
QY 121 tctcgttggagctctctcgcagctcgttaacccttaaccttagcttgatcactgctc 180
DB 121 TATCGGTGGAGCTCTTCGACGCTTGTAACTTAACCTTACCTTGTGATTCACCTGCTC 180
QY 181 gaagcaaacctcgaagagatgctctttagccttcgcgaggtttaacagatccggac 240
DB 181 GAAGCAAACTCCGAGAGAGTCTTTACGCTTTCGCGGAGTTAACAGATCCGGAC 240
QY 241 catgttccagagcttggatccaaactctgttaacctgttaacctgttccatgtccac 300
DB 241 CATGTTCTCCAGAGCTGGATCCAACTCTGTTAATCCTGTACCTGGTTCCATGTCAC 300
QY 301 tctaaccaagaacacgcgtcactcgtgtgatttggggaattcaacctctctgacat 360
DB 301 TGTAAACCAAGAACACCGCTCATCTGTTGATTTGGGGAAATCAAACTCTCTGGACAT 360
QY 361 ctctcgccttagcttgggaagcttgaacatttaccagatcttagagctctcaaaaacac 420
DB 361 CTTCGCGCTTAGCTTGGGAAGCTTGAACATTTACAGATCTTAGAGCTCTACAAAACAC 420
QY 421 atccaagaactatcacctcgaacttgaacatcgaagaatcaccatcagcttgcgtc 480
DB 421 ATCCAAGAACTATACCTCCGAACCTTGAAATCTGAAGATCTCATACAGCTTGATCAG 480
QY 481 tagaacaacaactcttaacaggatagctccacattcttgggaaatattgaagctctgctc 540
DB 481 TAGAACAACAATCTTACAGGATAGTTCACACTTCTTGGGAAATTTGAAGTCTCTGCTC 540

```

```

QY 541 ttttaacgcttaatgacacccgattgacgggccaatcccttagagcactcactgcaatc 600
DB 541 TTTTAAAGGCTTAATGACACCGATTGACGGGGCAATCCCTAGACACTCACTGCAATC 600
QY 601 ccaagccttaagttgttgaatgtctcaagaacatgatttgttggaaacatcccaaac 660
DB 601 CCAAGCCTTAAGTTGTGATGTCTCAAGCAATGATTTGTGTGAAACATCCCAAAAC 660
QY 661 ggaactttgctcaatccttcaagaacttgaagaacacccgagttggaaggacg 720
DB 661 GGACCTTTGCTCACAATTCCTTTACAGAACTTGAGAAACACCCGAGTGGAGGACCG 720
QY 721 gaattactgcttgaagctcgaacacacacacacacacacacacacacacacac 780
DB 721 GAATTACTCGGCTTGGCAACCTGACACTGACACTGACCTGAAAANAATTTGGCAAACTG 780
QY 781 aaatgaagaattgggggagaccttgaagaacacacacacacacacacacacacac 840
DB 781 AAATGAAGATTGGGGGAGACCTTGTAGAACACTTCACACTTATCAAAATATCACA 840
QY 841 tctactatgaataagatatatatagtcacacacacacacacacacacacacacacac 900
DB 841 TCTACTATGTAATAGTATATATATGATGATCCAAAAAATGAAGAAATCGAATCAGTA 900
QY 901 atatcatcgtctcaattgagaactttagagctctgtatgtaaaaattcttaaatgca 960
DB 901 ATATCATCTGCTCAATATGAACTTGAAGTCTGTGTATGTAATTTCTTAAATGCCA 960
QY 961 ctctcgccttagatgctcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 1020
DB 961 CTTCGCGTACTGTAATGTTGCGGTGTGGATCTCGAAGTAACTTGTATTTGTATG 1020
QY 1021 gtatcaagttgttgccttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 1063
DB 1021 GTATCAAGTTGTTCCTGCTGCTGCAAAAAAATTTGAAAAA 1063

```

```

RESULT 2
A67817 1106 bp DNA PAT 05-MAY-1999
LOCUS Sequence 22 from Patent WO9743427.
DEFINITION A67817
ACCESSION A67817
VERSION A67817.1 GI:4756639
KEYWORDS
SOURCE
ORGANISM
unclassified.
unclassified.
REFERENCE 1 (bases 1 to 1106)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICRITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source
location/Qualifiers
1..1106
/organism="unclassified"
/db_xref="taxon:32644"
142..798
/note="unnamed protein product"
/codon_start=-1
/protein_id="CAB42244.1"
/db_xref="GI:4756640"
/translation="MASRNYRWELFASLTLTALIHVLEANSSEDAIYALRSLTDP
DHVLSMDPTLVNPTCFWHTYCNODNRVTRDNLGNSLGHAPLGLKLEHLYLELY
KNNIQGITPSELGNLKNLISLDLYNNNLGIVPTSLGKLSLVFLRLNDRLTGPILR
HSRQSAFKVVDSNDLCGTIPINGPFAHIPLOFNENNPRLGPELGLASDYDINCT
"
BASE COUNT 331 a 258 c 206 g 311 t
ORIGIN

```

```

Query Match 85.8%; Score 912.4; DB 5; Length 1106;
Best Local Similarity 95.0%; Pred. No. 1.3e-224;

```


Db	557	TCITTTTAGGGCTTAATGACAAACCGATTGACGGGTCCAAATCCCTAGAGCACTACGGCA	956
Qy	559	tcceaagcccttaagctgtgtgtgacgtctcaagcaatgtatgtgtgtgacaacatcccaaca	658
Db	597	TCCCAAGCCCTTAAGATTGTTGACGCTCCAGCAATGATGTGTGGAACAATCCCAACA	656
Qy	659	acggacccctttgcacacatccctttaagaaacttggagaacaccggaattgtgaaggac	718
Db	657	ACGGACCCCTTTGCTCACATTCCTTTACGAACCTTTGGAAACCCAGATGTGAGGAC	716
Qy	719	cgaatctactcgctcttccaagctacgacacactactcagcttggaaaaattggcaaac	778
Db	717	CGGAATTACTCGGCTTGTCAGACGTACGACATTAATGCACTGAAACACTGGGAAAAC	776
Qy	779	tgaaaatgaagaatttgggggtgaccccttgaagaacacttcaaccactttatcaaatca	838
Db	777	TGAAAATGAAGAATTGGGGGTGACCTTGTAAGAACACTTCACACCTTTATCAAAATCA	836
Qy	839	catcacacatgaataagtatatatatgtagtcca--aaaaaaatggaagaatcgatc	896
Db	837	CATCATTAATGTAATAACTAATATATATGTAATAAAACAAAAAAATGAAGAATCGATC	896
Qy	897	agtaatatcatctggttcccaattgagaaccttggacttggcttgttatagtttaaattctaat	956
Db	897	GCTAATATCACTGCTGTCATGTGGAACCTTCGAGTTC--TGATGTAAAAATTCTTAAT	954
Qy	957	gcgaccttcgcgtaactgtaac	977
Db	955	CGGATTTTCGGCTTAATTAAT	975

```

RESULT      4
LOCUS       A67821
DEFINITION  A67821
ACCESSION   A67821
VERSION     A67821.1
KEYWORDS    GI:4756643
SOURCE      unidentified.
            unidentified
            unclassified.
REFERENCE   1 (bases 1 to 789)
            De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
            PRODUCTION OF APOMITIC SEED
            Patent: WO 9743427-A 20-NOV-1997;
            CIBA GEIGY AG (CH)
FEATURES             Location/Qualifiers
     source          1..789
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     <2..664
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAB42248.1"
                     /db_xref="GI:4756644"
                     /translation="RPTRPARYRWELPFAASLITLALIHVEANSEGDALVALRSRIT
                     DPDAILOSMDPTLVNCPWFHVTCTONDRYRVVDLGNNSGHLAPETGLKHLQYLE
                     LYKNNGISTIPSELGNKILSIDLYNNNNGIYPTSGIKLSIVFLRINDNRITGPI
                     PRLATATSLKVVVDYSSNDLCGTITPNCPRFAHITLQNFENNPRLEGPBLIGLASITDN
                     CT"
BASE COUNT      235 a      191 c      156 g      207 t
ORIGIN

```

Query Match	72.48;	Score 769.6;	DB 5;	Length 789;
Best Local Similarity	98.98;	Pred. No. 7.6e-188;		
Matches 775; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;

QY 103 caatcagcgctctcgaaactatcogtggagactcttcgcagcttcgtttaactcaacctta 152
| | | | |
Db 5 CCCACGCGTCCGCGAAATATTCGGTGGGAGCTCTTTCGAGCTTGTTAAATCTTAACCTTA 64
| | | | |
QY 163 gctttgatccaactcgtctcgaagcaaacctccgaagagatgctctcttcaagctcttcgcg 222

[illegible]

```

RESULT      5
A67823
LOCUS       A67823                894 bp    DNA
DEFINITION Sequence 28 from Patent WO97/43427.
ACCESSION   A67823
VERSION     A67823.1
KEYWORDS    GI:4756645
SOURCE      .
            unidentified.
            unidentified
ORGANISM    unidentified
            unclassified.
REFERENCE   1 (bases 1 to 894)
AUTHORS    De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE       PRODUCTION OF APOMITIC SEED
            Patent: WO 9743427-A 20-NOV-1997;
            CIBA GEIGY AG (CH)
FEATURES    (location/qualifiers)
            1..894
            /organism="unidentified"
            /db_xref="taxon:32644"
            <1..678
            /note="unnamed protein product"
CDS

```

[illegible]

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Features	Source	CDS
A67827	A67827.1	GI:4756649		thale cress.							
				Arabidopsis thaliana							
				Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.							
				1 (bases 1 to 2089)							
				De,V.S., Schmidt,E.D., Van,H.G. and Hecht,Y.F.							
				PRODUCTION OF APOMICETIC SEED							
				Patent: WO 9743427-A 20-NOV-1997;							
				CIBA GEIGY AG (CH)							
				Location/Qualifiers							
				1..2089							
				/organism="Arabidopsis thaliana"							
				/db_xref="taxon:3702"							
				/clone="SERK GENE CDNA"							
				195..2072							
				/note="unnamed protein product"							
				/codon_start=1							
				/protein_id="CAB42254.1"							
				/db_xref="GI:4756650"							
				/translation="MESSYHYFILLSTLLPNHSLMTASNLAEDELALTLRLTVLDPNN							
				NVLQSDPLVPCPTWFHNNMNSYIRVDLGAELSGHLYPELGIKNIQYEIYS							
				NITGPIPSNIGLNTLVSDIYINSGSPPEISGLSKIRPLRLNNNSLGTSPMS							
				LNITITLQVLDLSNNRISGSYPDNGSFLTPIFETANNLDGQVISHSPGSPSP							
				PEPTIPPPVSPSPSGVIGTGAIGAGVAGALPRAAALFAAMRRSPDIDFDPVA							
				EDEDEVLGOLKRFSLRELQVADSGFENKNIILGSGGKYKGLADGTVAVKRE							
				ERTGEGEIQETVEEMTSMVHRMLRLRGFCMTPTRLVYPYMAQSVASCLRRP							
				PSQPLPMPTRKRIALSGARGLSYLHPDCRIIHRDYKANILIDEEFEAVGDFGL							
				AKLMDYDHTVTAIRGIIHIAPEYISGSSSKTTPFGVIMLLELIGORAFDLA							
				PLADDVDMILMDWIKLEKKLEMLYDPLQCTYERRELDQVQVALLCTQSPMR							
				PKSEVYRMLEBGDLAEKWDKMOVVEILREETIDSPNDSMIIDSTINLHAEVLSGP							
				R"							
				BASE COUNT							
				1 569 a 419 c 502 g 599 t							
				ORIGIN							
				Query Match							
				Best Local Similarity 63.0%; Score 211.8; DB 5; Length 2089;							
				Matches 327; Conservative 0; Mismatches 192; Indels 0; Gaps 0;							
				184 gcaaacccgaaggagatgcctcttaacgctcttcgcgagatttaacagatccggaccat 243							
				270 GCTAATTGGAGAGCGATGAGCTTGGATACTTTGAGAGGTACTCTGATCCAAACAT 329							
				244 gtctccagagcttggatccaaactctttaaactcttgaacctggttcaagtcaactgt 303							
				330 GCTTCGACAGAGCTGGGAACTCAGCTAGTGAATCCTTCACATGGTCCATGTCACCTGC 389							
				304 aaccaagaacaacggcgtaactcgctgtgatttggggaattcaaacctctcggaaactt 363							
				390 AACCAACGAGACAGTGTATTAAGAGTTGATTTGGGGAATGACAGATTATCTGGCCATTTA 449							
				364 gcgccttgagcttgggaagcttgaacattacagtactagaagctctacaacaaacatc 423							
				450 GTTCCAGAGCTTGCGTGTCTCAAGAAATTGGCAGTATTGGAGCTTTACAGTAAACAATA 509							
				424 caaggaactatacccttcogaaacttggaaatctggaagaatctcaactgaagcttggatcgtaac 483							
				510 ACTGGCCGATTTCAGTATATTCTTGGAAATCGACAAACTTAGTAGTATGGATTTTAC 569							
				484 aacacaacattacagggaatagttcccaactctcttgggaaaaattgaagctctcgtcttt 543							
				570 TTAACACAGCTTCTCGGAGCTATCTCCGGAATCATTTGGGAAAGCTTTCAAGCTGAGATT 629							
				544 ttaagcttataagaacaccgatctgcggggccaatccctcctagaagcaactcaatgcaatcccca 603							
				630 CTCGCGCTTAAACAACAGTCTCACTAGGCTCAATCTCTAATGCTACACCAATTTTACT 689							
				604 acccttaaaagttctgaatctcacaagaaatgattgtcttgaacaaatccgaacaaagga 663							

```

Db 690 ACCCTTCAGTGTAGATCATCAATACAGACCTCTGGTGTACTTCCTCAATATGCG 749
Qy 664 cctttgtcaccattccttaccagaacttggagaacac 702
Db 750 TCCCTCTCACTCTTCACACCCATCAAGTTTGGCTAATAAC 788

RESULT 7
SB062279 936 bp mRNA PLN 09-DEC-1996
LOCUS Sorghum bicolor leucine-rich repeat-containing extracellular
DEFINITION glycoprotein mRNA, complete cds.
ACCESSION U62279.1 GI:1710123
KEYWORDS
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum.
REFERENCE 1 (bases 1 to 936)
AUTHORS Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.
TITLE Isolation of a cDNA encoding a novel leucine-rich repeat motif from
Sorghum bicolor inoculated with fungi
JOURNAL Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)
MEDLINE 9714217
REMARK Erratum: [[published erratum appears in Mol Plant Microbe Interact
1997 Mar;10(2):302]]
REFERENCE 2 (bases 1 to 936)
AUTHORS Hipskind,J.D.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant
Pathology, Purdue University, West Lafayette, IN 47907, USA
FEATURES
source
1..936
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone="SLRR"
12..755
/function="protein-protein ligand binding"
/note="leucine-rich repeat-containing extracellular
glycoprotein; contains six N-glycosylation sites
[IKS/T]"
/codon_start=1
/protein_id="AAC49559.1"
/db_xref="GI:1710124"
/translation="MAPOFAAGFTGLALATPASCNTBEDIYKORLAMEDPNVY
OSMNSTLANPCTWFVTCNNNNFVIRVDLGNAGISGPLPLAEIIONLYELGNGI
NGSIPETGNLTSLDMDNLFGELPTTGSYSTRYLKXONNLGTPIPSFGN
LTSLESTKQENSLSGAIPASIGNKALOFSLNDNMLGTVPKSPFLSTFGNLTEL
NTDRNNLDGTRTSSGLRYTAIIQDLAKTA"
36..77
/evidence="not_experimental"
BASE COUNT 247 a 232 c 207 g 250 t
ORIGIN

Query Match 14.6%; Score 155; DB 8; Length 936;
Best Local Similarity 58.2%; Pred. No. 1.2e-29;
Matches 272; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

Qy 187 aactcgaagagatgctcttaccgcttcgcgcgagtttaacagatcgcgagcatt 246
Db 81 AACACTGAAAGTGACATCTGTCAGAACGAAGGTTGGCATTGGAGGAGCAACCAACGCG 140
Qy 247 ctccagagctggagatccaactctgttaactctgttaccctgttccatgtccactgaac 306
Db 141 CTGAGAGCGTGAATTCGAGCTTGGCAATCCCTGCACCTGGTTCATGTCACCTGCAC 200
Qy 307 caagacaacgcgctcactcgtgtgatttgggggaattcaaacctctctggaactctgcg 366
Db 201 AACAAATAACTTTGTCAATCCGCGTGGATTGGGCAATGACAGGCAATCTCCGCTCTGCTT 260

```

```

Qy 367 cctgagcttgggaagcttgaacattacagtagtactagaagctctacaaaacacatccaa 426
Db 261 CCAAGATCTTCAGAAATTCAGAACTCCAGATACATGAGTGTATGGCAAGGCTTGAAC 320
Qy 427 ggaactatcctccgaacttgaactctgaacatcctacatcagcttgaatcgtgtacaac 486
Db 321 GGTTCGATTCAGAAACACCTAGGCAACCTGACATCTCATCGCTTGGATCTCTCGGAC 380
Qy 487 aacaacttaccaggagtagtccactcttgggaaattgaagtctctgtctttta 546
Db 381 AACCTCTTACCGCGCAAACTCCAACTACCTGTCTGTCTGTCAGCAGCTGCGATCTTG 440
Qy 547 cggcttaatgacacgcatgagcgggccaatccccaagacactcagatcccaagc 606
Db 441 AGGTGTACCAAAACACCTGACAGGCGCATATACATCATCTTTGGCAACCTGACTAGC 500
Qy 607 cttaagttgtgagtcacgaatgatttgttggagaacatccc 653
Db 501 CTTCTGGAATCGAAGCTTCAAGAGAAATCTGTAGCGGCGCTATTCC 547

RESULT 8
DCU93048 1755 bp mRNA PLN 28-JUN-1997
LOCUS Daucus carota somatic embryogenesis receptor-like kinase mRNA,
DEFINITION complete cds.
ACCESSION U93048
VERSION U93048.1 GI:2224910
KEYWORDS
SOURCE .
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
REFERENCE 1 (bases 1 to 1755)
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
TITLE A leucine-rich repeat containing receptor-like kinase marks somatic
plant cells competent to form embryos
JOURNAL Development 124 (10), 2049-2062 (1997)
MEDLINE 97313247
REFERENCE 2 (bases 1 to 1755)
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-1997) Molecular Biology, Agricultural University
of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
FEATURES
source
1..1755
/organism="Daucus carota"
/db_xref="taxon:4039"
34..1695
/note="SERK"
/codon_start=1
/product="somatic embryogenesis receptor-like kinase"
/protein_id="AAB61708.1"
/db_xref="GI:2224911"
/translation="MNRNSINILNTWQFDAYLDKYGVMITLEYISNNISGPIPSDIG
NLINIVSLDIYKNSFSGPIPDITLGIKILRFTPLRNNSISGPIPSLNIITITLYQVLD
SNRLSGVPVNDGNSFSLFTPIFANNLNLGCVTGPCGSPSPSPPTPSTVOP
PGONGPTGAJAGVAAAGALFLPAAPMAAMARRRPRHFDVAEDPEVHLQOLK
RFSRLQVATDFTSTILRGGFGYKORLADGSIYAVKRLKEBRTGGELQFOTEV
EMISMAVHRLRLRGFCMTPERLILVYVYMANGSVAAGLREOPSEPLDKPTKRI
ALGSAAGLSYTHDHPKTIHDDVYRANILDEEPVAVYDGLRIMDYKPTHTTA
VRCITLTIAPETISLTKSSSEKIDVRYGIMLELITIGQAPFLALANDDDVMLDWY
KSLTKERKLEMLVDPLENNYIDTEYQLIYVALCTQGSPEPERKMSVYMLBGDG
LAEKMDQKVEVTHQVLELAPRSEMLIDSTDLHAFLSGPR"
BASE COUNT 506 a 347 c 407 g 495 t
ORIGIN

Query Match 9.7%; Score 103.4; DB 8; Length 1755;
Best Local Similarity 55.4%; Pred. No. 2.4e-16;
Matches 200; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

```

[illegible]

QY	363	tggcgcttgagcttgggaagcttgaacatttaagtatcctcagaagctctcaaaaacaat	422
Dd	138	TGATGCTTACCTTGACAATAATGGGGTTCTTAAGACATTGGACCTTTACGCATAAACAT	197
QY	423	ccaaggaaactatacccttcgcgaacttgganaactctgaaaatctcatcgcttgaatctgta	482
Dd	198	AAGTGAACCAATTCCTAGTAGATCTTTGGGAATCGACAAAATTGGAGCTTGACCCTATA	25
QY	483	caaacacaactcttaagagtagtctccacctctttgggaaaaattgaagctctggtctt	542
Dd	258	CATGAATFACCTTCCTGGACCTATACCGGACAATTAGAAGAACTTACAAAGCCTAAGATT	317
QY	543	tttagagcttaatgacaacgatattgaagggccaatcccagaagcctcacatgcc	602
Dd	318	CTTGCGCTCCAAACAACAAGCCTCTCTGGTCCAATTTCCAAATGTCACTACTAATAATTATAC	377
QY	603	aagccttaaaagtctgtgatgtctcaagaacatgatttgtgtgnaacatccccaaaacgg	662
Dd	378	AACCTTCAAGTCGAGATTAATCAAACACGGCTATCGAGCACGATACCGGATTAATGG	433
QY	663	accctttgctcacatctcctttaacagaactttgagaacacacccgaggttggagagccgga	722
Dd	438	CTCATTTCTTTGTTTACACCTATCAGTTTTGCCAAAATTTCAATTATATGTGGACCCGT	497
QY	723	a 723	
Dd	498	A 498	

	RESULT	9		
	LOCUS	A67797		
	DEFINITION	A67797 .1815 bp DNA	PAT	05-MAY-1999
	ACCESSION	Sequence 2 from Patent WO/9743427.		
	VERSION	A67797		
	KEYWORDS	A67797.1 GI:4756623		
SOURCE		.		
ORGANISM		Daucus carota		
		Eukaryota; Vitidiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
		eudicots; Asteridae; euasterids II; Apiales; Apiaceles; Daucus.		
REFERENCE		1 (bases 1 to 1815)		
AUTHORS		De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.		
TITLE		PRODUCTION OF APOMICIC SEED		
JOURNAL		Patent: NO 9743427-A 20-NOV-1997;		
		CIBA GEIGY AG (CH)		
FEATURES		Location/Qualifiers		
SOURCE		1..1815		
		/organism="Daucus carota"		
		/db_xref="taxon:4039"		
		94..1755		
		/note="unnamed protein product"		
		/codon_start=1		
		/protein_id="CABA2241.1"		
		/db_xref="GI:4756624"		
		/translation="MNRSNINIILNYMOETDAYLDKYGVLMTIELYSNNISGP IPEDDLGL NLTIIVSLDIYNMNSRSGPIPTDLGRLTRLEPFLNMNSSLSGPIPMSITNTTIOVDIDL SNRRSGEPPDGSSFSLTFTPSFANNLNCIGPYTGRCBSPSPFPSTPTIPSVQZ PGONGRTGAAGCAAGAALFAAPAMAFMKMRKRKFREFPVPAEBDEVEHLSQL RSLTRLQVAIDTFESTILLGRGFERYVKRWLAGDSLVAVXRKEERTPEGELQTGLE ALESMNVARNHLRLRGFCMPCTERLLRVYPWANGSAASCIREOSPSEPILDWTPKKR ALSASAGLSYLHDHCDPKIIRRDYKAANIIMLEEFEAAVVDFGLARMDYKTHTVTI VRSTLCYIAPEYLGSKSEKTDVGVMGVMLLETIGORAFDLARANDDVALLDMW KLTKRKRELMVDDPLFNRY IDLPVEVOLIQVALICTOGSGPMWRPRKSYVMLEGGDC LAQRNDENQAEVIHQDELPHRTSERMIIDSINDTHAFLSGPR"		
BASE COUNT		531 a 354 c 415 g 515 t		
ORIGIN				

Query Match	9.7%;	Score 103.4;	DB 5;	Length 1815;
Best Local Similarity	55.4%;	Pred. No. 2.4e-16;		
Matches 200; conservative	0;	Mismatches 161;	Indels 0;	Gaps 0;

LEIRGENE	LOCUS	LEIRGENE	4604 bp	DNA	PLN	10-OCT-1996
DEFINITION	L.esculentum LRP gene.					
ACCESSION	X95269					
VERSION	X95269.1 gi:1619299					
KEYWORDS	LRP gene; LRP protein.					
SOURCE	tomato.					
ORGANISM	Lycopersicon esculentum					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; Solanales; Solanaceae; Solanum; Potatoes; Lycopersicon.					
REFERENCE	1 (bases 1 to 4604)					
AUTHORS	Vera, P.					
TITLE	Direct Submission					
JOURNAL	Submitted (18-JAN-1996) P. Vera, Universidad Politécnica de Valencia, Inst. de Biología Molecular y Celular de Plantas, Camino de Vera 14, E-46022 Valencia, SPAIN					
REFERENCE	2 (bases 1 to 4604)					
AUTHORS	Tornero, P., Mayda, E., Gomez, M.D., Canas, L., Conejero, V. and Vera, P.					
TITLE	Characterization of LRP, a leucine-rich repeat (LRP) protein from tomato plants that is processed during pathogenesis					
JOURNAL	Plant J. 10 (2), 315-330 (1996)					
MEDLINE	96367673					
FEATURES	Location/Qualifiers					
source	1..4604					
	/organism="Lycopersicon esculentum"					
	/catalytic="VFN8"					
	/db_xref="taxon:4081"					
	/tissue_type="leaf"					
exon	<1..345					
	/number=1					
gene	113..4355					
	/gene="LRP"					
CDS	join(113..345,2381..2452,2565..2708,2928..2999,3081..3152					

```

/product="LRR protein"
/protein_id="CAA64565.1"
/db_xref="GI:1619300"
/db_xref="SPTRMBL:Q96477"

```



```
misc_feature 1747..1838
/note="exon predicted by xgrall, quality excellent"
complement(2090..2250)
/misc_feature /note="exon predicted by xgrall, quality good_shadowexon"
2114..2246
/note="exon predicted by xgrall, quality excellent"
complement(2975..3024)
repeat_region /rpt_family="(GAA)n"
join(3070..3325,3565..3642,4094..4234,4338..4414,
4506..4655,4800..>5013)
/gene="F9f8.1"
3070..>5013
/gene="F9f8.1"
/note="similar to unknown protein GB:AAC27167 [Arabidopsis
thaliana]"
CDS join(3215..3325,3565..3642,4094..4234,4338..4414,
4506..4655,4800..5013)
/gene="F9f8.1"
/note="unknown protein"
/codon_start=1
/protein_id="AA01505.1"
/db_xref="GI:6016678"
/translation="MVGPAKQIVLFSSSIYVSGFHGNGAILSEVYARKADIIRG
YVGNSSRALEVDVQVPKDAVQPSLIVYFGNDSPHSSGLGPHVPLEYDVM
KIALHLQSDFTRIIFLSSPYDEAKVRQNSPYLSEVIRFNDLCKYSDACVLC
QELGLVYDLESTFOKADDMKTCFTDGIHLSAGSKIVAGEILRVYKEAMHPSLHM
KSMTEPADSPYDVSADKQVNSSEMTFWEEQMD"
complement(5513..5597)
repeat_region /rpt_family="(GA)n"
/protein_id="AA01508.1"
complement(5527..5579)
/misc_feature /note="exon predicted by xgrall, quality
excellent_shadowexon"
join(5589..5795,5897..5929,6013..6235,7264..7387,
7467..7727)
/gene="F9f8.2"
5897..7932
/gene="F9f8.2"
/note="similar to nucleic acid binding protein Alf1n-1
GB:AA20093 [medicago sativa]"
join(5696..5795,5897..5929,6013..6235,7264..7387,
7467..7727)
/gene="F9f8.2"
/codon_start=1
/product="putative nucleic acid binding protein"
/protein_id="AA01506.1"
/db_xref="GI:6016679"
/translation="MAAAVSSNPRTVEEIFKDYSAARALRLAKRDVDFSSQCD
EKENICLYGHNESMEVYNLPAEVPPLPPALGINFARQGMKRDMLSTVAYHSCW
LISVFTFGARLNNEKRRLSLINDPLTFDVYTGKAKNDKPPSSDSKSRNCTK
RSIDQTRSSTPKLMESYEEDDEEDHGDTLGS CGATYNEFWICDVCERYHG
KCVKITPAKASIKQYKCPCCAKKGRQ"
complement(7717..7760)
repeat_region /rpt_family="(CAT)n"
complement(7810..7847)
repeat_region /rpt_family="(GAAA)n"
8515..8572
/rpt_family="(TAA)n"
complement(join(<9426..9674,10198..10528,11098..11345,
11491..>11865))
/gene="F9f8.3"
complement(<9426..>11865)
/gene="F9f8.3"
/note="similar to leucoanthocyanidin dioxygenase
GB:BA420143 [perilla frutescens]"
complement(join(9426..9674,10198..10528,11098..11345,
11491..11865))
/gene="F9f8.3"
/codon_start=1
/product="putative leucoanthocyanidin dioxygenase"
/protein_id="AA01507.1"
/db_xref="GI:6016680"
/translation="MNNIDKIKSKTCLNDQEQEVKIDNMHMSDDCKNRIKNSG
LGERWPEPIYVAVOSIAESNLTSLPDRYIKPPSQRPQTITIDHQPEVADINIPDIIDS
```

```
LFSGNEDDKRISACRNEGFQYINHGKPELMDAARETKSFPLPEAEVYSNS
PRTEGGSRIGVEKGAIDWMDYIYHLEPLAKPFNKPSPISIRMDNDEYKEL
VKIAGRLMTILISSNLGRAQIQEAGGDEYVACLVNYYPKQPELIGLSPSDP
GGMTILPDDQVVLQVRHGDWTIVNPLNHAIVIGQOILSSKSKSYEHVIV
NSERKSVLAFFYVNPKSDIPIQMOQLVTSMPPLPMTFDQRLFIHQPRKSH
VESHSRPR"
repeat_region /rpt_family="(GA)n"
11962..11996
/misc_feature /note="exon predicted by xgrall, quality
marginal_shadowexon"
complement(13367..13468)
/misc_feature /note="exon predicted by xgrall, quality marginal"
complement(13742..13788)
/misc_feature /note="exon predicted by xgrall, quality marginal"
complement(join(13822..14179,14261..14398,14502..14582,
14676..14861,14941..15033,15124..15190,15272..15361,
15662..16380))
/gene="F9f8.4"
complement(13822..16380)
/gene="F9f8.4"
/note="identical to omega-3 fatty acid desaturase,
chloroplast precursor GB:P46310 [Chloroplast Arabidopsis
thaliana]"
CDS complement(join(13988..14179,14261..14398,14502..14582,
14676..14861,14941..15033,15124..15190,15272..15361,
15662..16355))
/gene="F9f8.4"
/codon_start=1
/product="Omega-3 fatty acid desaturase, chloroplast
precursor"
/protein_id="AA01508.1"
/db_xref="GI:6016681"
/translation="MAWDVYSEGGIRPLPIRTTPPSNLTSSNNKRPRLSSSYKTS
SSPLSGUNSRDGFTRNVALNVSITPLTPIFEESPLEDNRKQFDPGAPPPNLDIR
AIPRKHQVKNPKMSLSYVVDVAIVFALAGAAIYNMIVPLYLAAQGTWMALEF
LGHDCGSGSESNPKLNSVVGHLHSLILVPYGMRIISHRTNHNHGYENDESMFPM
SEKIVNLDKRTFRFFRTLLPYMLAVPEVLMANSPKSGSHYHPSDLPLEPKQVL
```

Query Match 8.0%; Score 84.8; DB 8; Length 101284;
Best Local Similarity 54.98; Pred. No. 1,7e-11;
Matches 167; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```
QY 350 tctcgaacatctcgagctgagctggaagcttggaacattacagtagagctc 409
||| ||| || || || || || || || || || || || || || || || ||
Db 44936 TCTTGGTCAAGTTCATCTTCGATTCGATTCGATTCGATTCGATTCGATTC 44877

QY 410 acaaaacaaacatcaaggaactacacctcgaactcgaactcgaactcgaatc 469
||| ||| || || || || || || || || || || || || || || || ||
Db 44876 CTTCACACAAATTCCTTGCTGTAATCCCATCTCTTTGGCAATCTAACCCAGCTGATCG 44817

QY 470 gcttgagctcgaacaaacattcagagatagctccactcttgggaaatga 529
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44816 TCTTACAGGTTATTCCTCAATAGCTCAGTGAAGATGTTCCATCTCACTGCAATTTGA 44757

QY 530 agctctggtctcttaacgcttaatgacaacagattgacggggccaactagagcac 589
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44756 CCAAGGTGCTCTCTTATTACTTCCACAACTCAGTGAAGATGTTCCATCTCACTGCAATTTGA 44697

QY 590 taatgcaatcccaagccttaagtgtgtgtagtctcaagaatgattgtgtggaaca 649
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44696 TCAAGTTACTATCAACATTGATGACTTTGAAGCAAGTAAACAGCTTCACTGGAATC 44637

QY 650 tccc 653
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44636 TCCC 44633
```

RESULT 12
AC008017 116944 bp DNA PLN 18-SEP-1999
LOCUS AC008017 Arabidopsis thaliana chromosome I BAC F3N23 genomic sequence,
DEFINITION complete sequence.

ACCESSION	AC008017
VERSION	AC008017.2
KEYWORDS	GI:5487890
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE	1 (bases 1 to 116944)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Chen,S., Harman,P., Hicks,R., Huerta,M., Mason,S., Siepel,J., Zimmerman,M., Buehler,E., Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Vyotskaia,V.S., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 116944)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Chen,S., Harman,P., Hicks,R., Huerta,M., Mason,S., Siepel,J., Zimmerman,M., Buehler,E., Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Submitted (10-JUN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	3 (bases 1 to 116944)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Chen,S., Harman,P., Hicks,R., Huerta,M., Mason,S., Siepel,J., Zimmerman,M., Buehler,E., Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Submitted (15-JUL-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
TITLE	4 (bases 1 to 116944)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Brooks,S., Buehler,E., Chao,Q., Dunn,P., Gonzalez,A., Khan,S., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Submitted (18-SEP-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
TITLE	On Jul 15, 1999 this sequence version replaced gi:5441915.
COMMENT	location/Qualifiers
FEATURES	1..116944
source	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="F3N23"
	complement(1333..2966)
gene	/gene="F3N23.1"
	complement(join(1333..2271,3355..2966))
CDS	/note="81% identical to threonine synthase [Arabidopsis thaliana] (gi 4914408). Location of ests GBG589 3' (gb f15267) and GBG615 (gb 226034)."
	/codon_start=1
	/product="Putative threonine synthase"
	/protein_id="A055628.1"
	/db_xref="GI:5903070"
	/translation="MASFLPHATYFPSSHSETSLKPSAASFTVCTASAPVPPQT POKRSPDENIDEARRRHOLQNSARVPPNAPPSSTESYDEIYRSOSGLL DQHDRAALKRVDGEFMRNLFDRNGTKTPMGSSGVSKSEWVLPIDDDIDVSAFEG NSMLFMRERKQYLDQNDLMAVHCGISHTSGFKDLSVLYSQVNRLLKRNKPFYIGV
gene	6030..7079
	/gene="F3N23.2"
	6030..7079
CDS	/note="Unknown protein: Location of ests 205D127 (gb H7203) and 205D12P 3' (gb AA605559)."
	/codon_start=1
	/protein_id="A055629.1"
	/db_xref="GI:5903071"
	/translation="MTGAAEESAEQETLPAIDWEMDKSFEVIGALFSGVSA LPFAVLKTRQOVCHSQSGCITATLVNBEELRGVIRGFGISLNGTIPARLVYAL EYKNSGVSAVSLGLEAKAAVANAAGLSAAADLVMTVPVAVASQRLMVGASG LVNASCNVNGDFAFRKIVRADPGKLGFGISILTYAPSNVAVMASVVAQRMW GGIGCYVCKDEESGNNSTYMKDPSKTIAMAVOGVAALIGVSALITPDLTKTRIQ VLDGEESNNKRGPSIGQTVNRNLVREGWTAQYRGDGRCA SMSATMTITVEFL KRLSAKNHGFYSK5"
	complement(8272..9661)
gene	/gene="F3N23.3"
	complement(join(8272..8658,8746..8916,9005..9079,9161..9252,9364..9661))
CDS	/gene="F3N23.3"
	/note="99% identical to Transcription factor [Arabidopsis thaliana] (gi 2398525)."
	/codon_start=1
	/product="Transcription Factor"
	/protein_id="A055630.1"
	/db_xref="GI:5903072"
	/translation="MHOMLNKKSATHTSLPYNLTISGVVPTDVSANRGSASEL STVDSRPHGHTTKQISFODSSSTSGSYEVAASGDDNSROISFAKSGSE ITRKRGASHPKQSGSTGPNITPAPAQNFSHYADHPYGLIATYTPQAPTCNPO MVAIPGRVPLPAELTETDPEVNAKQYIMRQRQAKLAKNKLARKPFIHES RHYHALKRPKRGSGRFLNKLQSESOAAARQEDKLGQOVNRTNNSREAHNLQ NNRDSSTSGSDITSVDQADIFGHTERFQSGFPPIINRAMLVHGSNDMHGDDMH FFSVHT"
	complement(12076..16057)
gene	/gene="F3N23.4"
	complement(join(12076..12519,12613..13851,13955..14140,14368..15451,15672..16057))
CDS	/gene="F3N23.4"
	/note="Similar to disease resistance proteins Rpp1-Wsb (gi 3860165). Rpp1-Wsc (gi 3860167) and Rpp1-Wsa (gi 3860163) [Arabidopsis thaliana]."
	/codon_start=1
	/product="Similar to disease resistance proteins"
	/protein_id="A055631.1"
	/translation="MASSSSSATRLRHVDVLFSPFGVDRQRTIVSHLYVALRNNGVL TKDDKRLKTEGDTLADGIYKAIQTQSWPAVYVILSENAVSTMCLEELRLMOJHSEOI KVDFITGVKPSDYRIOEGSFATAPQSVDEADDAIAYVGGISRLRPMKSTDLINYG MEAHMKMTLLNIGCEDEVHMTGIMWGGIGSTAKCLIDYRFSQFPAHCFLEWVS KGYDIKHLQKLELSHLIDEDVELMSHAGSQNIKRLHOKRVVLDVADVEDQIAG LAKDPQSWFGSGRIIITTRDGLNCSGVNIIEVCLDDKALQVFKKLADGGRDPS DGFEOLEFIRASRLAHGCPALVAFASLSAIVADIMEDDEALLETFPKKNQVILRA SYGDIQDYDKTVFLHVAQCFRNGHILRYIRAFELNCARINHLAAKLQVINSIDGICM HILLYVOTGREIVROESMRPSKQFLMDPFEIHYVYDSNTHLGNVSNQILSDVYL SRNLKLMWAVPTLIPFLFRPHITIELSRISKLNSLMDGKLPNLRILIDVYQS NLRLEPDLSTAVNPELIELSCSTLVQIPISIRKLRLKLNMMYCGELGCVILNDLQ EASLWSRGLKRIILNPHSGATLSLTDLIAQKIFIKLSGSGTDHLSFSVQVTA HOSVTHLNSGFEGLKSLDKRFSYRLDPSYRDFCSFADPCLTEKLINTNEIDTPE DICOLQITLIDGGNDFYVLPISMGQIALKXISNCRKALPOLSOVERLHNSK CVKLSGLMGLTAGRYVLDPCYCKGSLDSLNGIISVEKSAQRNLELISLENKS LVSLSELSHTKLTLYDLSLFRRLPYSIRLSFMRLLYLNCKIRISLIDPESL KYLVAGCESLEHVNFSNHSFNHLDPSHCSIECISDVRDMNEYSOAEFRLVC IITKYSIASITNMTSRWEPRIKLPKIKAPKLVGFVQVIMVVCERPFHLOPAPFYIN WDCQSGEDELITLKNPNLYOSESSENNNNPYYKWHLVIVQIPTGISLSEIDVQCES HIOVPEGRDEITLLCGVEHVGFLK"
	16543..17907

SVNREIETIGKVRHNLKLEEFWLRKREXGLILTYMENGSLHDIHEFTNPKPLDMS
TRHNIATVGTAGHAIYHFDCAIYHRIKPNAILDLDELPHIDFIAKLIDQSAT
SIPSTVQCTIGIAPENAFITVKRSSESVYGVLLLETRKALDPSNGEDIV
GWSRVWOTGEIOLKIVDPSLDELIDSSVMEQVTEALSLARCAEKAEVDRKPTMDV
VNOLTFWMSRSYSSSVRNRSK"

BASE COUNT 1365 a 1060 c 1017 g 1591 t
ORIGIN

Query Match 7.7%; Score 81.4; DB 8; Length 5033;
Best Local Similarity 49.9%; Pred. No. 1.1e-10;
Matches 205; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 249 ccagagctggagatccaaactctgttaactctgttacctgtccatgtaaccca 308
DB 1281 CCAAGAGCTGGAGATCCCAATTCACCTCTGTCTTCAAGCTGGAGATGAGCA 1340
QY 309 agacacacggcgctacactgctgtggttggtgggaattcaaacctctctggacatctggcc 368
DB 1341 GAGACAAATTTGTTATCTGTAACCTCTCTCTATGAGATCTCAGGCAATTCGGGCC 1400
QY 369 tgaagctggagagcttgaaattacagatctagagcttacaacaaacacacccaagg 428
DB 1401 CGAATATCTCGCATTTGAGCAATTTGAGAGAGTGTCTCAGTGGCAATGTTCTTGG 1460
QY 429 aacatccttcggaacttgaaatctgaagaatctcatcagcttgatctgacacaa 488
DB 1461 CTCATATCTCCAGATGAGCAATTCAGCTCTTGTGAACACATGATGTGCTCCAA 1520
QY 489 caattcagaagagatgctccactctcttggaagaattgaagctctgctcttctttag 548
DB 1521 CAGCTTACTGTAATNCCCTGACACCTTGAGCTTTCAGAAATTAAGAATTAG 1580
QY 549 gcttaagacacacgatgagagggccaatccctagagacactcaatgcccaagcct 608
DB 1581 CCTGTCTTAAATCTCTGATGGCCCATTCCTGAGTCTTCAATTCACATTT 1640
QY 609 taaagcttgatgctcgaacatgatttggtgagacacacccaacaa 659
DB 1641 AGAAACGTTTAACTTCACTGCAATGCTTAAATGTTCAATCCTTCANA 1691

RESULT 14
AC021198 81513 bp DNA PLN 09-FEB-2000
LOCUS Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete
DEFINITION
AC021198
VERSION Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete
AC021198.2 GI:6957696
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE
AUTHORS
1 (bases 1 to 81513)
Liu, S.-X., Sakano, H., Yu, G., Lee, J.-M., Lenz, C., Pham, P., Toriumi, M.,
Chin, C., Chlou, J., Choi, E., Chung, M., Gonzalez, A., Hwang, B.,
Liu, A., Vayenberg, M., Altai, H., Brooks, S., Buehler, E., Cho, O.,
Conn, L., Conway, A. B., Hansen, N. F., Johnson-Hopson, C., Khan, S.,
Klin, C., Lam, B., Miranda, M., Nguyen, M., Palm, C. J., Shinn, P.,
Southwick, A., Davis, R. W., Ecker, J. R., Federspiel, N. A. and
Theologis, A.
Arabidopsis thaliana chromosome 1 BAC F14D7 sequence
Unpublished
2 (bases 1 to 81513)
Theologis, A.
REFERENCE
AUTHORS
Direct Submission
Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
TITLE
JOURNAL
3 (bases 1 to 81513)
Theologis, A.

TITLE
JOURNAL
Direct Submission
Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT
The sequence is of BAC F14D7 from Arabidopsis thaliana chromosome
1. The sequence does not represent the sequence of the entire
insert of this clone. It is shorter by 6954 bp because we submit
only the unique sequence of the clone. However, in order to
facilitate the joining of overlapping clones in the future for
creation of larger contigs, we provide small overlaps (200 bp)
between overlapping submitted clones. The 5' end of this sequence
overlaps by 200 bp to the 3' end of the sequence of the clone
F1504.

FEATURES
Source Location/Qualifiers
1. 81513
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F14D7"

BASE COUNT 27779 a 13814 c 13868 g 26052 t
ORIGIN

Query Match 7.5%; Score 79.8; DB 50; Length 81513;
Best Local Similarity 53.7%; Pred. No. 3.2e-10;
Matches 165; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 347 accctctggaacatcttgccgctgagcttggaagcttgaaacttaacagatcagagc 406
DB 2466 ATCTACAGTGGGATTCGCCCGAATAGGCAACATGATGATGATGATGAGT 2525
QY 407 tctacaaaacacatccaaagacatctacatccggaactggaactggaactggaactc 466
DB 2526 TGAGTACACAACTTACCGGTTCAATCTCTTCTTACGAAATCTAAGAAATTTGA 2585
QY 467 tcaagcttgatctgacacacacacatcttaacagagatgctccactcttggaat 526
DB 2586 CTATCTTACTTACTTACGAGAACTATTTAACTGCTGTCATTCGCCCGAATGAGCA 2645
QY 527 tgaagctctgctcttttaagcttaagcaacacgattgaaggggccaatccctag 586
DB 2646 TGAATTCATGATTTGATTTACGTTGAATTAACAACAACTTACCGGTTCAATCTTCT 2705
QY 587 caatcaccgaatcccaagccttaagcttgatgctcgaagaatgattgtgaa 646
DB 2706 CCTTGGAAATCTCAAGAACTTGAACATTTATCTTACCTGAATTTTAACTGCTG 2765
QY 647 caatccc 653
DB 2766 TCATTC 2772

RESULT 15
ATAC011765 134402 bp DNA PLN 08-DEC-1999
LOCUS Arabidopsis thaliana chromosome 1 BAC F1M20 genomic sequence,
DEFINITION
ATAC011765
VERSION Arabidopsis thaliana chromosome 1 BAC F1M20 genomic sequence,
ATAC011765.4 GI:6539234
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE
AUTHORS
1 (bases 1 to 134402)
Lin, X., Kaul, S., Town, C. D., Benito, M., Creasy, T. H., Haas, B.,
Ronnig, C. M., Koo, H., Fujii, C. Y., Utterback, T. R., Barnstead, M. E.,
Bowman, C. L., White, O., Niernman, W. C. and Fraser, C. M.
Arabidopsis thaliana chromosome 1 BAC F1M20 genomic sequence
Unpublished

REFERENCE	2 (bases 1 to 134402)				
AUTHORS	Lin, X. and Kaul, S.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-OCT-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org				
REFERENCE	3 (bases 1 to 134402)				
AUTHORS	Lin, X.				
TITLE	Direct Submission				
JOURNAL	Submitted (08-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA				
COMMENT	On Dec 8, 1999 this sequence version replaced gi:6102640. Address all correspondence to: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA e-mail: xlin@tigr.org BAC clone F1M20 is from Arabidopsis chromosome I and is near the molecular marker m1425. The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of three methods: Gene prediction programs including GRL (available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), Genscan (Chris Burge, http://genome.stanford.edu/~chris/genscan.html), and NcapPlantGene (http://www.cbs.dtu.dk/netpene/chnspene.html), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (http://www.tigr.org/tcdb/at.html). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by RepeatMasker (Arian Smil, http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRL are annotated as misc features.				
FEATURES	Location/Qualifiers				
Source	1..134402				
	/organism="Arabidopsis thaliana"				
	/cultivar="Columbia"				
	/db_xref="taxon:3702"				
	/chromosome="I"				
	/map="m1425"				
	/clone="F1M20"				
misc_feature	138..269				
	/note="exon predicted by xgrail, quality marginal_shadowexon"				
misc_feature	complement(229..281)				
	/note="exon predicted by xgrail, quality excellent"				
misc_feature	complement(393..453)				
	/note="exon predicted by xgrail, quality excellent"				
misc_feature	complement(744..1286)				
	/note="exon predicted by xgrail, quality excellent"				
misc_feature	1271..1354				
	/note="exon predicted by xgrail, quality good_shadowexon"				
misc_feature	complement(1356..1475)				
	/note="exon predicted by xgrail, quality excellent"				
misc_feature	1378..1581				
	/note="exon predicted by xgrail, quality good_shadowexon"				
misc_feature	complement(1509..1530)				
	/note="exon predicted by xgrail, quality marginal"				
misc_feature	complement(join(<1655..1777,1908..2135,2258..2581,2679..2963,3052..3429))				
	/gene="F1M20.1"				
gene	complement(<1655..>3429)				
	/gene="F1M20.1"				
	/note="similar to p56 protein kinase GB:AAB59449 [Homo sapiens]; contains Pfam profile: PF00069 Eukaryotic protein kinase domain				
	complement(join(1655..1777,1908..2135,2258..2581,2679..2963,3052..3429))				
	/gene="F1M20.1"				
	/codon_start=1				
	/product="putative protein kinase"				
	/protein_id="AF15905.1"				
	/db_xref="GI:6539235"				
	/translation="MGCVSKQTVSVTPAIDHSGVFKDNEKSGSGRIYEDPPRP LKTIYSRSGRSQKRSOKSGSEIGSGASDVSFRLGNVRYLAEQVAAGPAWL SVNAGPAIGHWPLRSDAPFEKLEKIGQGTYSNPAVAYETGRIYALKRYAFDEE SVKEMARETILRLNHPILIKELGISKISCNLOLYEFEMEDIDGLSSPIKKT TPIIKCYMKQLISGLDCHSRGMARDIKGSNLLSNGLIKVADPCLANSGNSGK KPLISRYVTLMTRPPELIGATIDGASVDMSGCTFAELLRPLIRGRTYEQQL KIFKICGSPEDYMKRSKLPKPAHLKRPQYTDMSKIFELKDLSEITNLIFETLSLP HKRGTSASALVSQVFTYTRFADPSPSLIYPPSKRIDIKRHDARSVISFT"				
	complement(4069..4131)				
	/note="exon predicted by xgrail, quality excellent"				
misc_feature	complement(join(<4608..4760,4885..>4974))				
	/gene="F1M20.2"				
mrna	complement(<4608..>4974)				
	/gene="F1M20.2"				
gene	complement(<4608..>4974)				
	/gene="F1M20.2"				
	/note="similar to dolichyl-phosphate mannosyltransferase polypeptide 2 GB:4503365 [Homo sapiens] (regulator of dolichol phosphate-mannose synthesis: EMBD J 1998 Sep 1:17(17):4920-9)"				
	complement(join(4608..4760,4885..4974))				
	/gene="F1M20.2"				
	/codon_start=1				
	/product="putative dolichyl-phosphate mannosyltransferase polypeptide 2"				
	/protein_id="AF15906.1"				
	/db_xref="GI:6539236"				
	/translation="MELADRAVGLLSSISIFNYTFWVILPEVDSDFIKTKYL PQYKALVYPRAGIALISISYFIGKVMKSKKKA"				
	complement(5111..5190)				
	/note="exon predicted by xgrail, quality excellent"				
misc_feature	complement(<5342..>7603)				
	/gene="F1M20.3"				
mrna	complement(<5342..>7603)				
	/gene="F1M20.3"				
gene	/note="contains Pfam profile: PF01348 Type II intron maturase"				
	complement(5342..7603)				
	/gene="F1M20.3"				
	/codon_start=1				
	/product="putative type II intron maturase"				
	/protein_id="AF15907.1"				
	/db_xref="GI:6539237"				
	/translation="MAKRETMFSLAGELASIVESSSHVDDSKPRSRMLKRSLEL LKRRKEQDINGFSDLLKVIYARETLRDYDIDRLNSNVSTIERGSAFDSIAE ELSSGVFVASNTFSIVARDKTEVLYLPSVALKVOEIRIVLEVPSPRFSIS CRSRGASALAKYINNINSISRDWCFILSNKLIDVSEENLISVEEVESSSILL RSMFEARVNLIEFGPGKGHGPOECSVRLVNIYLDREDFEYIRSRHAGLDS KTDSDSGSKLNSMFRROAGEOGKIDETODVAFSGIRVCTYRKRYRSPVRYHKKK VRPALOKEPATLTGTVRIGKKRLGGLKAKVESEETKGADSNSTLSISCKRAGK TDMWYKTLRTIMEDVLTSTADSEEFVLSKVEEPTVPEQLRTADGARYSNMILDS SETANLEALPPOSHDRPVTFGGDVAPNAGRLRYGLITADGARYSNMILDS TAOIIDVYSGLVRRWVYIYEGCSNPEIDALIDNOIRMGCTITLAKRIHENEIERK LDLELSTIPAEIDIOEHKLDSPAFDRDHLTYGNSGCLGLSTARLYSESRPC NCFYAGSMAPAVYTLHAMERQKRFPGKMTGVSVCIPSSLNRRATGLCKHLDNYIG QISQANDPGAMP"				
	complement(7885..8010)				
	/note="exon predicted by xgrail, quality marginal"				
misc_feature	8771..8821				
	/note="exon predicted by xgrail, quality marginal_shadowexon"				
misc_feature	8935..9008				
	/note="exon predicted by xgrail, quality marginal_shadowexon"				
repeat_region	9421..9528				

```
misc_feature      /rpt_family="POLY_A"
9791..9912
/note="exon predicted by xgrail, quality marginal"
misc_feature      /note="exon predicted by xgrail, quality good"
10574..10662
<10992..>14231
/gene="FLM20.4"
<10992..>14231
/gene="FLM20.4"
gene             /note="similar to brassinosteroid insensitive 1
GB:AC49810 (putative receptor protein kinase); contains
Pfam profiles: PF00560 Leucine Rich Repeat (17 repeats),
PF00069 Eukaryotic protein kinase domain"
10992..14231
CDS              /gene="FLM20.4"
                /codon_start=1
                /product="putative receptor protein kinase"
                /protein_id="AF15908.1"
                /db_xref="GI:6539238"
                /translation="MAAIVAGDSLDSREVLISLKSYLESRNPONRGLYTEMKMENQ
DYVQMPGICTPQSRVRVGINITDSTISGPLFKNSALTELYIDLNRTEGEIPD
DLRCHNLKHLNLSHNLBGEISLPGLSNLEVLIDLNLKRTIGDIQSPLFCNSLYVA
NLSTNNEVGRIDIDFNGCRNLKYVDFSSNRFGEVWTFGRVDFSVADNHLSGNTISA
SMFRGNTLQMLDLSGNAFGEFPGQVSNCONLVNLNKNFTGNIPAEIGSISLK
GLYGNNTFSRDIPELTNLNLNVLFLDSLNRKFGDIOEIFGRFTQKYLVLHANSYV
GGINSSNIIKLPLNLSRLDLGNNESGOLPTEISOISLKEFLIAYNNFSGDIPQEXGN
MPLQALDLSFNRLTGSIPASPKLTSLIMTMANNSISGEIPREIGNCTSLMENVVA
NNQISGRFHELTMRGNSNPPTPEVNRQNRDKIINSGECLAKRKWIPADPEPPENVY
AIIKRSKRSRLMDVHLKGLFPCVSGSTVRLKISATIDLSGNKSGEIPASISQW
DRLLTLHGFNEFEGKLPPEIGQLPLAFNLTRNNSGEIPQELGNLKLQNLDSLQN
NFGNNEPTSLNDLNLSEKFNISYNPFSIGALPTTGQVAFYDKDSFLGNPLRPPSFN
OSGNNEKISNOVYGNRPRTLLIWLISALAFIACLVSGIVLVKASREAEIDL
LDGSKTRHDMTSSSGSPWLSGKIKYIRLDSFTYADILKATNSSEERYVGRGY
GTYVRGLPDRGRYAVKKLQREGTEAEKPRAMVILSANRQDMAHPNLYLQWCL
DSEKILVHEIMGGSLBELITDTKRLQMKRKIDINLVARGLYFLHHCYRSIVHRD
VKASNVLLDRHGNAKVDFGLARLNLVNGDSHVSFVLAGTIGYVAPETGQWQATTIKGD
```

```
Query Match      7.5%: Score 79.2; DB 8; Length 134402;
Best Local Similarity 52.8%: Pred. No. 4,7e-10;
Matches 171; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
```

```
QY 329 tggatttgggaatcaaacctctcaggacatcttgcgctgagcttggagaacttgaac 388
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12043 TAGATTGGCGCTACACAACTTCTCGGAGCAGTTACTACTGAATTTCTCAGATACAGA 12102

QY 389 attcaagatcctagagctctacaacaacaacatccagaactatcctccgaacttg 448
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | |
DB 12103 GTTGAAGTTCTTGATTCTTGCTTATTAATACTTCAGTGGCATATACACAGAGATATG 12162

QY 449 gaaatctgaagaatctcatcagcttgatctgtacacaacaacatcttcaaggatagttc 508
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12163 GGAACATGGCGGGGCTTCAGACACTTGATCTCTCTTAAACAAGCTGACCGTTGAGATAC 12222

QY 509 ccactctcttgggaaatgaagtctctgctcttlltcaggcttaatgaacacgatga 568
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12223 CAGCTTCATTGGGAATTCCTCTTTGTGGCTAATGCTTGCACAAACACTCTCTAT 12282

QY 569 cggggccaatccctagagcaactcactgcaatcccaagccttaagtgtgtatgtctcaa 628
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12283 CAGGAGAAATCCCTCGAAGATGGCAACTGACAAAGCCTTTTGCGTTAACGTGCGCA 12342

QY 629 gcaatgatttgttggaaacatcc 652
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12343 ACAACCAAGCTCTGTGTAGATTCC 12366
```

Search completed: June 24, 2000, 00:26:34
Job time: 46549 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:53:35 ; Search time 446.21 Seconds
(without alignments) 596.029 Million cell updates/sec

Title: US-09-180-798-30

Perfect score: 1063
Sequence: 1 tcgaccacgcgtccgcagc.....tcgcaaaaaaaaaaaaaaaaaa 1063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1063	100.0	1063	1 V06590	Arabidopsis thaliana
2	912.4	85.8	1106	1 V06586	Arabidopsis thaliana
3	841.8	79.2	981	1 V06587	Arabidopsis thaliana
4	769.6	72.4	788	1 V06588	Arabidopsis thaliana
5	361	34.0	894	1 V06589	Arabidopsis thaliana
6	211.8	19.9	2089	1 V06591	Arabidopsis thaliana
7	103.4	9.7	1814	1 V06571	Daucus carota SERK
8	75.4	7.1	3176	1 T62124	Arabidopsis thaliana
9	75	7.1	5940	1 X23526	O. longistaminata X
10	71.4	6.7	3573	1 T06307	Partial tomato pat
11	71.4	6.7	6471	1 T06306	Tomato pathogen re
12	70.8	6.7	4081	1 V06585	Arabidopsis thaliana
13	70.8	6.7	6695	1 V06570	Daucus carota SERK
14	70.6	6.6	3979	1 V14518	CF-5 pathogen resi
15	70.6	6.6	3979	1 V14519	CF-5 pathogen resi
16	70.6	6.6	4123	1 V14523	CF-5 pathogen resi
17	70.2	6.6	9424	1 X23525	O. sativa Xa21 gen
18	67.8	6.4	3541	1 V14522	CF-5 pathogen resi
19	66	6.2	8416	1 X23523	O. longistaminata
20	65.4	6.2	19639	1 X23524	O. longistaminata
21	63.6	6.0	3921	1 T31300	Rice Xa21 disease
22	63.6	6.0	13340	1 X23522	O. longistaminata
23	61	5.7	3293	1 X23532	Tomato Xa21 clone
24	60.4	5.7	7204	1 X23527	O. longistaminata X
25	59.2	5.6	3045	1 X23531	Maize Xa21 gene DT
26	58	5.5	2075	1 T49435	Tomato polygalactu
27	56.4	5.3	2192	1 X23530	Maize Xa21 gene DT
28	55.8	5.2	1554	1 T31307	Tomato RRR gene cl
29	55.8	5.2	4104	1 X07356	Arabidopsis sterol
30	55.4	5.2	1058	1 T49434	Pear polygalactu
31	55.4	5.2	6256	1 T31329	Rice Xa21 disease
32	53.2	5.0	3842	1 X23533	Tomato Xa21 clone
33	53	5.0	5733	1 X00477	Arabidopsis thaliana
34	49.8	4.7	3050	1 T06509	Tomato pathogen re

35	49.8	4.7	3089	1 T47877	Tomato pathogen re
36	48.4	4.6	2880	1 Q91450	Tomato Cf-9 cDNA.
37	48.4	4.6	2880	1 T06304	Tomato pathogen re
38	48.4	4.6	3905	1 Q91449	Tomato Cf-9 gene.
39	48.4	4.6	3905	1 T06303	Tomato pathogen re
40	41.4	3.9	1117	1 Q42595	Sequence encoding
41	41.4	3.9	2917	1 T86755	Raspberry dtul gen
42	40.6	3.8	792	1 Q42596	Sequence of varian
43	39.4	3.7	1686	1 N10014	Sequence encoding
44	37	3.5	692	1 X13578	Enterococcus faeca
45	36.4	3.4	2241	1 T91078	Staphylococcus aur

ALIGNMENTS

RESULT 1
ID V06590 standard; cDNA to mRNA; 1063 BP.
AC V06590;
DT 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS plant breeding; leucine-rich repeat; ss.
KW Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 106..762
FT /tag=a
FT /note="shows high homology to SERK"
FT
FT W09743427-A1.
PN 20-NOV-1997.
PD 13-NOV-1997; E02443.
PF 14-MAY-1996; GR-010044.
PR (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI; 98-086529/08.
DR P-PSDB; W47022.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 86-88; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;

Query Match 100.0%; Score 1063; DB 1; Length 1063;
Best local similarity 100.0%; Pred. No. 3.6e-287;
Matches 1063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	tcgaccacgcgtccgcagcgaacccaatttgcctcctcaatcgttcgcgaataact 60
DB	1	TCGACCCACGCGTCGACGAAACCAATTTGCTTCCATCTTGTCAGAAATTTACT 60
QY	61	caattcctttagatctctctccttcgcacctcgcatagcacatgscgtcgaac 120
DB	61	CAATTCTCTTAGATTACTCTCTTCGACCTCGCATGCTCAGATGGCTTGGAAC 120
QY	121	tatcgtgtagcgtccttcgcagcttcgttaaccttaaccttagcttgatcactg 180
DB	121	TATCGGGGAGGAGCTTCGACAGCTTGTAATCTTAACCTTAGCTTGATCCTGGTC 180
QY	181	gaacgaacctcgaagaagatgctcttcgcgtccttcgcgcgaagttaacagatcgcgac 240
DB	181	GAAACAACCTCCAGAGAGATGCTTTAGCGCTCTTCGCCGAGTTTAACATCCGAC 240
QY	241	catcttcagaagcttggaatccaactctgttaacctgttaacctgttacctggtcacc 300
DB	241	CATCTTCCAGAGCTGGAGTCAACTCTTGTAACCTTGTAACCTGGTTCATGTCACC 300
QY	301	tgtacaaagacacacgcgtcactcgtgtggaatttggaataacactcttgaat 360
DB	301	TGTAAACAAGACACCGCTCACTCGTGGATTGGGAATTCAACCTCTTGACAT 360
QY	361	cttcgcctgagcttgggaagcttgacattacagtlactagtagctctacaacaaacac 420

```
Db 361 |||||CTTGCGCTTAAGCTGGGAAAGCTTGAACATTTACAGTATACAGCTTCACAAAAAACAC 420
Qy 421 atccaaagaaactaactcctcgaacttgaaatctgaagaatcactcagcttgatcgt 480
Db 421 ATCCAAAGAACTATACCTTCGGAAGTGAATCTGAAAGTATCATACAGCTTGATCTG 480
Qy 481 tacaacaacaactcttaagagatagttcccaactcttttggaataattgaagtctcgtgc 540
Db 481 TACAACAACAATCTTAAGGAGATAGTCCACCTTTGGGAAATGGAAGTCTTGCTG 540
Qy 541 ttttaagccttaagcaacacgattgaagggccaactccctagagcaactcactgaatc 600
Db 541 TTTTAAAGGCTTATGACAAACGATTGACGGGGCAATCCCTAGAGACTCATGCAATC 600
Qy 601 ccaagccttaagttgttgatgtctcaagaatgattgtgtggaacaatcccaacaac 660
Db 601 CCAAGCCTTAAAGTTGTGATGTCTCAAGCAATGATTGTGTGGAACAATCCCAACAAC 660
Qy 661 ggaacctttgtccatcctcttcaagaactttgagaaacccgaggttgagagaccg 720
Db 661 GGACCTTTGTGCTACATTCCTTTACAGAACTTGAGAAACCCGAGGTTGAGGGACG 720
Qy 721 gaactactcgtcttgcaagctacgacactaactgcacctgaaanaattggcaaaactg 780
Db 721 GAATTACTCGGCTTGCAGAGTACGACACTGACACTGAAANAATTTGGCAAAACCTG 780
Qy 781 aaatgaagaatcgggggggtgacctgtgaagaacactcccaacttataaataacaca 840
Db 781 AAAATGAGAAATTTGGGGGTGACCTTGTAGAACACTTCACCACTTATCAATATACCA 840
Qy 841 tctactatgtaataagatataatataatgtaagtcacaaanaaaatgaagaatcgaatcga 900
Db 841 TCTACTATGTAATAGATATATATATATGTAATGCCAAAAAATGAAATCGAATCGTA 900
Qy 901 atatcactcgtctcgaacttgagaacttgagctgtgtatgttaaaatcttaaatgcga 960
Db 901 ATATCATCTGCTCAATTTGGAATTTGAGACTTGTGAGCTGTGTAAATTTCTTAATCGCA 960
Qy 961 ctctcgcgtactgtaaatgctgtgtgtgtggagttctgagaagtaacatttgatgtgatg 1020
Db 961 CTTTCGCTACTGTAAATGTTGTTGGTGTGGATCTGAGAAATTAACATTTGATGTGATG 1020
Qy 1021 gtatcaaatgtctcgtcctgtctgcacaaaaaataaaaaa 1063
Db 1021 GTATCAAACTTGTCTGCTCTCTGCAAAAAAATAAAAAA 1063

RESULT 2
ID V06586 standard; cdna to mRNA; 1106 BP.
AC V06586;
DN 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; epomixis; apomictic; seeds; production; embryos;
RM plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT 142..798
FT /*tag= a
FT /*note= "shows high homology to SERK"
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;
DR P-PSDB; W47018.
PS Production of apomictic seeds - useful in plant breeding
CC Claim 28: Pages 71-73; 123pp: English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
```

```
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;
Query Match 85.8%; Score 912.4; DB 1; Length 1106;
Best Local Similarity 95.0%; Pred. No. 3,6e-245;
Matches 1011; Conservative 0; Mismatches 36; Indels 17; Gaps 6;
Qy 14 ccgaagaaaccttaattgtcttcctatc-tgtgtcagaanaat-----actca 62
Db 39 CCAAGCTAACCTTAAATTTTGTGCTTCTCTCTCTTTTGAAGAAATTTCCCTTACTCTCA 98
Qy 63 aattcctatagatctctctctcctcgaactcgaatagctacatgagctcgcgaacta 122
Db 99 AATCTCTTTGATTTCCCTCTTTAAACCTCGAAGAGTCACATGCGCTCGAAACTA 158
Qy 123 tcggtggagagctctcgcgaactcgttaactccttaacttgatttgattaccggtgcga 182
Db 159 TCGGTGGAGAGCTCTTCGAGCTTGTTAACCTTAACCTTACCTTTGATTCACCTGTGCA 218
Qy 183 agcaaacctcgaagagatgctcttcttaagccttcgcgcgaggttaacaagaatccgaca 242
Db 219 ACCAAATCTCGAAGAGATGCTCTCTACGCTCTTCGCGAGTTGACAGATCCAGACCA 278
Qy 243 tgtctcagaagcttgatcccaactctgttaactccttgtaacctggttccatgtacactg 302
Db 279 TGTCTCCAGAGACTGGATCCCAACTGTTAATCTGTACCTGTACCTGTCATGTCACTG 338
Qy 303 taaccaagaacacgcgcacactcgtgtgatttgaggaaatcgaacctcctctgacaact 362
Db 339 TACCAAGCAACCGCGTACTCGTGTGATTTGGGAATTCAAACCTCTGTGACATCT 398
Qy 363 tgcgcctgagcttgaggaaacttgaaacattacagatcctagaagctctacaaaacaaat 422
Db 399 TCGGCTGAGCTTGGAAGCTTGAACATTTACAGATCTAGAGCTTACAAACCAACAT 458
Qy 423 ccaaggaactaactcctcgaacttgaaatcctgaagaatctcaactcagcttgatcgtga 482
Db 459 CCAAGGAATATATACCTTCGAACTTGAAATCTGAAGATCTCATCACTTGATGCTGTA 518
Qy 483 caacaacaactctaaagagatagttcccaactccttttggaanaattgaaagctcgtgctt 542
Db 519 CAAACAATCTTTAAGGAGATGTTCCCACTTTCTTGGAANAATTGAAGTCTGAGCTT 578
Qy 543 tttaagccttaatgaacacgattgacgggggccaatccctcctagagcaactcgaatccc 602
Db 579 TTTAGGCTTAATGACAACCGATTCACCGGTCCAAAT-CTTAGAGACTGACAGCAATGCC 637
Qy 603 aagcc-ttaaagttgtatgtctcaagcaatgattgtgtggaacaatcccaacaacg 661
Db 638 AAGCCTTTAAAGTTGTGACGCTCAACGCAATGTTGTGGGACAATCCCAACAAACG 697
Qy 662 gacctttgtcacattccttcaagaacttgagaacacccgaggttgaggagaccg 721
Db 698 GACCTTTCTACATCTCTTACGAACTTTGANAACAACCCGAGATTGAGAGGACCGG 757
Qy 722 aattactcgtcttgaaactcgaactcgaactcgaactgcaaaaaattggcaaaactgga 781
Db 758 AATTACTCGCTTGCAACCTAGCACTACCTACCTGCAAACTGGAACCAAACTCGA 817
Qy 782 aaatgaagaattgggggggacctgtgaagaacacttaccacttatcaaatcacat 841
Db 818 AATGGAAGAAATTTGGGGGAGCTGTGAAGAACACTTCAACCTTTATCAATATCACAT 877
Qy 842 ctactatgaataaagatatatactagtcgaa--aaaaaaatgaagaatcgaatcagt 899
Db 878 CTATATGTAATAGATATATATATGTAATGTAATGTAATGTAATGTAATGTAATGTA 937
Qy 900 aataatactgctcgaacttgagaacttgaggtctgtgatgtgaatctcgaatcga 959
Db 938 AATATCATCTGCTCAATTTGAGAACTTGCAGGTC--TGTATGTAATTTCTTAATCG 995
Qy 960 acttcgcgtactgtaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1019
```


ID	Accession	Gene	Location/Qualifiers	Score	DB	Length
Db	996	ATTTCGCTACTGTAAAGTTCGGTTGGGATTCGTGAGAGTAACATTTGATTGTAT		79.2%	81.8	981
Qy	1020	gatacaagttgtctgccttgctctgcaaaaaa		96.1%	1.7e-225	1063
Db	1056	GGTATCAAGTTGTCTGCTTGTGCAAAAAA		0	32	1099
<p>Query Match: 79.2% Score 81.8 DB 1 Length 981; Best Local Similarity 96.1%; Pred. No. 1.7e-225; Matches 885; Conservative 0; Mismatches 32; Indels 4; Gaps 2;</p>						
Qy	59	ctcaattcctttagtactctctcttcgacccctgagtagctacatgagctctgca		79.2%	81.8	118
Db	57	CTCAATTCTTTCGATTTCCCTCTTAAACCTCCGAAGCTCATGCGCTCTCGAA		96.1%	1.7e-225	116
Qy	119	actatcggtggagctctcgcagcttgtaacccaacttagcttgattacacctg		79.2%	81.8	178
Db	117	ACTATCGGTGGAGCTCTCGCAGCTTCTTAACCTTAACCTTAGCTTGATTCACCTGG		96.1%	1.7e-225	176
Qy	179	tggaggaacaccccggaagagatgctcttaagctcttcgacggaggttaacagatccg		79.2%	81.8	238
Db	177	TGGAAGCAACATCCGAAGAGAGATGCTCTTACGCTCTTTCGCGAATTGACAGATCCAG		96.1%	1.7e-225	236
Qy	239	accatgcttcagagctgagatcccaactctgttaactccttgtaactggtccatgta		79.2%	81.8	298
Db	237	ACCATGCTCTCCAGACCTGGATCCCAACCTTGTTAACTCTTGATCTGTTCCATGTCA		96.1%	1.7e-225	296
Qy	299	cctgttaacaaagaacccgcgtcactcgtgtgatttggggaattcaaacctctctgac		79.2%	81.8	358
Db	297	CCGTGAACCAAGCAACCCGCGTCACTCGTGTGGATTTGGGAATTCAACTCTCGGAC		96.1%	1.7e-225	356
Qy	359	atttcgctgagcttgggaagcttgaacattacagatcatagaagcttcaaaaaa		79.2%	81.8	418
Db	357	ATTTCGCTGAGCTTGGAAGCTTGAAACATTTACAGTATCTAGAGCTCTACAAAAA		96.1%	1.7e-225	416
Qy	419	acatcaaggaactataccttcggaacttggaaattgagaatctcatcagcttgatc		79.2%	81.8	478
Db	417	ACATTCACAAAGAAATTAACCTTCGAACTTGGAATCTGAAGATCTCACTACGTTGGATC		96.1%	1.7e-225	476
Qy	479	tgtacaacaacaactcttaagagatagtcctccactcttctgggaattggaagctctg		79.2%	81.8	538
Db	477	TGTACAACAACAATCTTACAGGATAGTTCCTCTTGGGAAATTTGAAGTCTCTG		96.1%	1.7e-225	536
Qy	539	tcttttacggttaatgaacacgattgacgggccaatcccttagagcaacttaactgca		79.2%	81.8	598

QY	Db	537	TCITTTTACGCGCTTAATGACAAACGATTGACCGGCTCCAAATCCCTAGACAGACTCAGCGCAA	596
QY	599	tcccaagccttaagaatgltgtga	lgtctcaagaacatgattgtgtggaacaatccacaa	658
Db	597	TCCCAACCTTAAAGTGTGTTGACGTCTCAACGAAATGATTTGTGGAACAATCCAAACA		656
QY	659	acggaccttgttcaatcttcc	taacgaactttgagaacaaccgaggttggaggaac	718
Db	657	ACGACACCTTTGCTCAATTCCTTTACAGAACTTTGGAACAACCCAGATTTGAGGAGAC		716
QY	719	cggattaccgcgtcttcgaagctacgaagcaactatgcacttgaataaatgtgcaaac		778
Db	717	CGGATTTACCGGCTCTGCAACCTACGACACTAAGTACTGCACTGAAACAACTGGCAAAAC		776
QY	779	tgaaatgaagaattgagggtgagccttgtaagaacacttcaaccattacaataca		838
Db	777	TGAAAAAGAAAGAAATTTGGGGGGTGACCTTGTAAAGAACTTACACCACTTTATCAAAATATA		836
QY	839	cattactatgataataagatatataatgagtcaca--aaaaaaatgaagaatgcgaac		896
Db	837	CATTATTAATGATTAATAGATTAATATATAGTACTAAAACAAAAAATGAAGAAATCGAATC		896
QY	897	agtaatacatcgtgtcgaattggaacttggaggtcgtgtatgtgtaataattccta		956
Db	897	GGTATATATCATCTGTCCTCAATTGGAAGAACTCGAAGTC--TGATATGAATATTTCTAAT		954
QY	957	gcgaccttgcgcgtacgtgaat		977
Db	955	GCGATTTTCGCTTAATTAATCT		975
RESULT	4			
VO6588	ID	VO6588	standard: CDNA to mRNA; 788 BP.	
AC	VO6588;	03-AUG-1998	(first entry)	
DT	Arabidopsis thaliana	SERR	LRR homologous EST clone.	
DE	receptor kinase; apomixis; apomictic; seeds; production; embryos;			
KM	plant breeding; leucine-rich repeat; ss.			
KW	Arabidopsis thaliana.			
OS	Location/Qualifiers			
PH	Key			
FI	CDS			
FT		2..664		
FT		/*tag=	a	
PN		/note=	"shows high homology to SERR"	
PR	WO9743427-A1.			
PD	20-NOV-1997.			
PR	13-MAY-1997: E02443.			
PR	14-MAY-1996: GB-010044.			
PA	(NOVS.) NOVARTIS AG.			
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;			
DR	PI WPJ; 98-08629//08.			
DR	P-PSDB; W47020.			
PS	Production of apomictic seeds - useful in plant breeding			
CC	Claim 28: Pages 79-80; 123pp: English.			
CC	The sequence is that of an EST clone showing high homology to			
CC	SERR LRR (leucine-rich repeat) sequences.			
SQ	Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;			
Query Match	72.4%;	Score 769.6;	DB 1;	Length 788;
Best Local Similarity	98.9%;	Pred. No. 2.1e-205;		
Matches	775;	Conservative 0;	Mismatches 9;	Indels 0;
			Gaps	0;
QY	103	caatgagcgtctcgaacatactcgtgtggaagctcttcgacgttcgttaactcaactta		162
Db	5	CCCAAGCGCTCGGGAACATCATCGTGGGAGGCTTCGACGCTGTGTAATCCTAACCTTA		64
QY	163	gcttgattaccgtgtcgaagaacactccgagaagatgctcttaacgcttcgcgg		222
Db	65	GCTTGATTAACCTGCTCGAAGCAACTCCGAAGAGATGCTCTTTACGCTTCGCCG		124
QY	223	agttacaacatccgagacatgltcttcgaagcttggatccaactctgttaactcttgt		282

```
Db 125 AGTTTAACGATCCGACCATGTCTCCAGAGCTGGATCCAACTCTTTTAATCCTTGT 184
Qy 283 accgtgtccatgctcaactcgttaacaaacacacgcgcgtcactcgtgtggaattgggaat 342
    |||
Db 185 ACCGTGTTCCATGTCACCTGTAACCAACACCGCGTCACCTGCTGATTTGGGGAAAT 244
Qy 343 tcaaacctctcggaaattcttcgcctcgtcgttggaaattgaaacatttaacatctta 402
    |||
Db 245 TCAAAACCTCTGGACATCTTGCGCTGAGCTTGGAACTTGAACATTTAAGATCTTA 304
Qy 403 gagcttcaaaaaaacacataccaagaataataccttcgaaacttgaaatctgaagaat 462
    |||
Db 305 GAGCTTACAAAACAAACATCCAGAGAACTATACCTTCGAACTTGGAAATCTGAAGAAAT 364
Qy 463 ctcatcagcttgatctgtacaacaaacaaatcttaacaggaatagttccactctcttggga 522
    |||
Db 365 CTCATAGCTTGGATCTGTACAAACAACATCTTACAGGATAGTCCCACTCTTTGGGA 424
Qy 523 aaattgaagtctcgtgtcttttaaggcttaattgacacacgaattgacgggacaaatccct 582
    |||
Db 425 AAATTGAAGTCTGTGCTTTTACGGCTTAATGACACCGAATGACGGGGCAATCCCT 484
Qy 583 agagcactcactgcaatcccaagccttaagtggtgagtctcaagcaatgatttgt 642
    |||
Db 485 AGAGCACTACCTGCAATCCCAAGCCTTAAAGTGTGATGTCTCAAGCAATGATTTGTGT 544
Qy 643 ggaacaatcccaacaaacggaccttttgcacatctcttaacagaactttgagaacaaac 702
    |||
Db 545 GGAACATCCCAACAAACGAGACCTTTTGTACATTCCTTTACAGAACTTTGAGAACAAAC 604
Qy 703 cgaaggttggaggagacggaatctcgtctcgaagctacagacatacgaacactga 762
    |||
Db 605 CCAAGGTGGAGGAGACCGGAATTAATCTGCTTCAAGCTACAGCACTAATGCACTGA 664
Qy 763 aaaaattggcaaaacctgaaatgaaagaaattgggggtgaccttgaagaacacttcacc 822
    |||
Db 665 AAAAATTGGCAAAACCTGAAATGAAGAATTTGGGGGTGACCTTTGAAACAACTTCACC 724
Qy 823 accttcaaatatcatcatctactatgtaataagatatataatggtgtccaaaaaaata 882
    |||
Db 725 ACTTTAATCAATATCAATCTACTATGTAATATATATATATATATATATATATATATAT 784
Qy 883 tga 886
    ||
Db 785 AAAA 788

RESULT 5
V06589
ID V06589 standard; cDNA to mRNA; 894 BP.
AC V06589;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS 1..678
    /tag= a
    /note= "shows high homology to SERK"
PT WO9743427-A1.
PD 20-NOV-1997.
PE 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR P-PSDB; W47021.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 83-84; 123pp; English
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;
```

```
Query Match 34.0%; Score 361; DB 1; Length 894;
Best Local Similarity 72.8%; Pred. No. 2e-91;
Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2;

Qy 189 ctccgaagagatgctcttcaacgtcttcgcccggagcttaacagatccggaacttct 248
    |||
Db 15 CTCGAAAGGGAGCGCTCTTACCGCGCTTGCCGAGACTATACGATCCAGACAAATGTTGT 74
Qy 249 ccagagctggatccaaactctgttaatcctgtacctgggtccatgtaacaa 308
    |||
Db 75 TCAGAGTTGGATCCAACTCTGTATATCCTGTACTTGTTCAAGTCACTGATATCA 134
Qy 309 agaaacccggtcactcgtgtgatttgggaattaaacctctcggacatcttcgccc 368
    |||
Db 135 ACACCAATCAAGTCACTGCTGTGATTTGGGAATCAAACTATCTGACATCTATGATAC 194
Qy 369 tgaacttgggaagcttgaacattacagatc----- 402
    |||
Db 195 TGAACCTGGGAAGCTTGAACATTTACAAATATCTGTATGAAATCATCATCTTTGCTTT 254
Qy 402 ----- 402
Db 255 TGATTATCTGAAAACATTTACATTATCATCATATATATATATATATATATATATATAT 314
Qy 402 --- agagctcacaacaaacacatccaagaactataccttcgaacttggaaatctgaa 458
    |||
Db 315 TAGTGAACCTCAAAAACGAGATTCAGGAACTATACCTTGTAGCTTGGAAATCTGAA 374
Qy 459 gaatcactcagcttggatctgtacaacaaacatcttaacaggaatagttccactcttt 518
    |||
Db 375 GAGCTATATAGTGTGGATCTGTACAAACAAATCTACACGGGAAATCCCACTCTTT 434
Qy 519 gggaaattgaaagtcctgtgtcttcttaacgcttaacagacacggatgacgggccaat 578
    |||
Db 435 GGGAAATTTGAAGTCACTGTTGTTTGGCGCTTAAAGAAACCGATGACCGGCTCTAT 494
Qy 579 cccctagagcactcactcgaatcccaagccttaagtggtgagtctcgaacatgatatt 638
    |||
Db 495 TCCTAGAAACTCAACAGTATTTTAACCTTAAAGTTGTAAGTGTCAAGGAATGATTT 554
Qy 639 gtgtgaaacaatcccaacaaacggacctttgtcacaactccttlaacagaacttgaagaa 698
    |||
Db 555 GTGTGAACATTTCCAGTAGAAGACCTTTGAACACATTCATATGCAAAACTTTGAA 614
Qy 699 caaccggagcttggaggagacccgaataactcgcgtctcgaagctaaagacatacgcac 758
    |||
Db 615 CAACCTGAGATTGGAGGACCGAAGCACTAGAGTCTTGCAGAGCTATGACACCAATTCAC 674
Qy 759 ctgaaaaaattggcaaaacctgaaatgaaagaaattgggggtgaccttgaagaacactt 818
    |||
Db 675 TTTAAAGAAAGTGAAGAACTATATTAAGAAAGAA-TGTTAGTGACCTTTGAAGAACTCTG 733
Qy 819 caacccttatcaaatatc 837
    |||
Db 734 TACCAAGTGTGTATATC 752

RESULT 6
V06591
ID V06591 standard; cDNA to mRNA; 2089 BP.
AC V06591;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK gene.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS 195..2072
    /tag= a
    /product= SERK protein
PT WO9743427-A1.
```

PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GS-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: M47023.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 21; Pages 91-95; 123pp; English.
CC The sequence is that encoding SERK, a putative receptor kinase.
CC It may be used as part of a method of producing apomictic seeds
CC comprising: (a) transforming plant material with a nucleotide
CC sequence encoding a protein which in active form in a cell or
CC cell membrane renders the cell embryogenic; (b) regenerating
CC the transformed material into plants or carpel-containing
CC plant parts; and (c) expressing the sequence in the vicinity
CC of the embryo sac. The apomictic seeds and embryos thus produced
CC can be developed into plant progeny. This is useful in plant
CC breeding programs. Controllable and reproducible apomixis provides
CC many advantages in plant improvement and cultivar development in
CC the case that sexual plants are available as crosses with the
CC apomictic plant. Apomixis provides for true-breeding, seed
CC propagated hybrids and could shorten and simplify the breeding
CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;

Query Match 19.9%; Score 211.8; DB 1; Length 2089;
Best Local Similarity 63.0%; Pred. No. 1.1e-49;
Matches 327; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 184 gcaaacctcgaagagatgctcttaccgtctccgcgaggttaacagatccgagcaat 243
DB 270 GCTAATTGGAAGGTGATGCTTGCACATCTTGAAGGTTACTTACTTATCCAAACAT 329
QY 244 gtctcagagagcttgagatcaactctgttaactctgttaccgttccatgctcactgt 303
DB 330 GTCTTGACAGAGCTGGAGTCTTACGCTAGTGAATCTTGACACATGTTCCATGTCACTTGC 389
QY 304 aaccaagaacaaccgcgtcaactctgtgagatttgggaatcaaacctctctgacatct 363
DB 390 AACAAAGAGAAAGCTGTCAATAGATTGATGGGAAATCGAAGTATCTGGCATTTTA 449
QY 364 gcgcctgagcttgggaagcttgaacatttcaagtatctgagagcttcaaaaaaacatcc 423
DB 450 GTTCCAGAGCTGGTGTGCTCAAGAAATTTGCAATTTGAGAGCTTTCAGATAACAATTA 509
QY 424 caaagaactatacttcgaacttgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 483
DB 510 ACTGGCCCGAATCTCTACTAATCTTGGAATCTGACAAACTTACTGATTTGGATCTTAC 569
QY 484 aacaacaatttcaagagatgagttccacttcttgggaataatltgaagttctgtgctt 543
DB 570 TTAACACAGCTTCTCCGGTCTTATCCGGAATCAATTTGGGAAGCTTTCAAAGCTGACATTT 629
QY 544 ttacggcttaatgacaacgatttgaaggggcaatcccttgaagcattcgaatccca 603
DB 630 CTCGGGCTTAACAAACAGTCTACATGGGCTCAATTCCTATGTCATGACCAATATTAAT 689
QY 604 agacctaaagtgtgtatgtctcaagcaatgatttgttggaaacaatcccaacaacgga 663
DB 690 AACCTTCAAGTGTAGATCTATCAATAACAGACTCTCTGTTCAGTCTCTGACATGCG 749
QY 664 cctttgtcaccatcttcaagaacttgaagaacac 702
DB 750 TCCTTCTACTCTTCAACACCATCATGTTTGTGATTAAC 788

RESULT 7

V06571
ID V06571 standard; cDNA; 1814 BP.
AC V06571;
DT 03-AUG-1998 (first entry)
DE Daucus carota SERK gene.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW plant breeding; ss.
OS Daucus carota.
FH Key Location/Qualifiers
FT CDS 94..1755
FT /tag- a
FT /product- SERK protein
FN W09743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GS-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: M47013.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 21; Pages 47-51; 123pp; English.
CC The sequence is that encoding SERK, a putative receptor kinase.
CC It may be used as part of a method of producing apomictic seeds
CC comprising: (a) transforming plant material with a nucleotide
CC sequence encoding a protein which in active form in a cell or
CC cell membrane renders the cell embryogenic; (b) regenerating
CC the transformed material into plants or carpel-containing
CC plant parts; and (c) expressing the sequence in the vicinity
CC of the embryo sac. The apomictic seeds and embryos thus produced
CC can be developed into plant progeny. This is useful in plant
CC breeding programs. Controllable and reproducible apomixis provides
CC many advantages in plant improvement and cultivar development in
CC the case that sexual plants are available as crosses with the
CC apomictic plant. Apomixis provides for true-breeding, seed
CC propagated hybrids and could shorten and simplify the breeding
CC process so that selfing and progeny testing to produce and/or
CC stabilise a desirable gene combination could be eliminated.
CC Apomixis allows plant breeders to develop cultivars with
CC specific stable traits for such characteristics as height,
CC seed and forage quality and maturity.
SQ Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

Query Match 9.7%; Score 103.4; DB 1; Length 1814;
Best Local Similarity 55.4%; Pred. No. 1.8e-19;
Matches 200; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 363 tgcgcctgagcttgggaagcttgaacatttaccagtatctagagctctacaacaacat 422
DB 138 TGATGCTTACCTTGACCAAAATATGGGGTCTTATGACATTGGAGCTTTCACGCAATTAACAT 197
QY 423 ccaaggaactatacttcgaacttgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 482
DB 198 AAGTGGACCAATTCCTAGTATCTTGGGAATCTGCAAAATTTGGTGAAGTGTGACCTATA 257
QY 483 caaacaacatttcaagagatgttccacttcttgggaataatltgaagttctgtgctt 542
DB 258 CATGAATAGCTTCTCTGACCTATACCGGACACATTAGGAAGCTTTCACAAAGCTTAAGATT 317
QY 543 ttacggcttaatgacaacgatttgaaggggcaatcccttgaagcattcgaatccca 602
DB 318 CTTCGCTCTCAACAAACAGCCTCTCTGCTCAATTTCCAAATGCTACTGCTAATATTAATAC 377
QY 603 aagacctaaagtgtgtatgtctcaagcaatgatttgttggaaacaatcccaacaacgga 662
DB 378 AACCTTCAAGTGTAGATCTATCAATAACAGACTCTCTGTTCAGTCTCTGACATGCG 437
QY 663 acccttgcacatcttcaagaacttgaagaacacccgaggttggaggagacgga 722
DB 438 CTCATTTCTTGTGTTTACACCTATACGTTTGGCAATAATTTGAATTTATGTGACCCGT 497
QY 723 a 723

Db 498 A 498

RESULT 8

ID T62124 standard: cDNA to mRNA; 3176 BP.

AC T62124;

AT 10-JUN-1997 (first entry)

DE Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.

KM Plant; morphogenesis; regulation; short; stem; alteration;

KW Inflorescence; extraneous; gene; expression; transformation;

KM Increase; control; form; length; ds.

OS Arabidopsis thaliana.

FH Key Location/Qualifiers

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

DT 17-JUN-1999 (first entry)

DE O. longistaminata Xa21 gene family member A2 DNA.

KM Xa21; receptor kinase-like protein; multigene family; RRR; rice;

KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.

OS Oryza longistaminata.

PN WO9909151-A2.

PD 25-FEB-1999

PF 17-JUL-1998; U14841.

PR 13-AUG-1997; US-910386.

PA (RSCG) UNIT CALIFORNIA.

PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,

PI Wang G;

DR WPI: 99-204431/17.

PT New RRR polynucleotides and nucleic acid constructs - used for

PS generating transgenic plants resistant to Xanthomonas

PS Claim 1; Page 52-53; 6/PP; English.

CC This invention describes a method for conferring disease resistance in

CC plants. The invention describes the use of novel genes and proteins

CC belonging to the Oryza longistaminata and Oryza sativa receptor

CC kinase-like protein (RRK) Xa21 multigene family. Such genes from

CC cassava, maize and tomato are also described. The genes and proteins can

CC be used for enhancing resistance to Xanthomonas in a plant, preferably in

CC rice or tomato.

SQ Sequence 5940 BP; 1570 A; 1200 C; 1188 G; 1982 T;

Query Match

Best Local Similarity 7.1%; Score 75; DB 1; Length 5940;

Matches 165; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Db 343 tcaaacctctctgacacatctgagcttgagcttgagacattacagatctca 402

Db 2406 tccgacatctgctgggagatctgctgctgggagacatctgctgctgctg 2465

Qy 403 ggcctctcaaaaacacatccaaggaactacccctgagcttgagaaatcga 462

Db 2466 gacctcagcagacacacacacacacacacacacacacacacacacac 2525

Qy 463 ctcaacagcttgatctctacacacacacacacacacacacacacacacac 522

Db 2526 ctccac 2585

Qy 523 aactgaagctctgctctctctctctctctctctctctctctctctctct 582

Db 2586 aatcctac 2645

Qy 583 agagac 2705

Db 2646 tcatctctgagacacacacacacacacacacacacacacacacacac 2765

Qy 643 ggaac 2720

Db 2706 agttccatctctctca 2720

Query Match

Best Local Similarity 7.1%; Score 75; DB 1; Length 5940;

Matches 165; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Db 343 tcaaacctctctgacacatctgagcttgagcttgagacattacagatctca 402

Db 2406 tccgacatctgctgggagatctgctgctgggagacatctgctgctgctg 2465

Qy 403 ggcctctcaaaaacacatccaaggaactacccctgagcttgagaaatcga 462

Db 2466 gacctcagcagacacacacacacacacacacacacacacacacacac 2525

Qy 463 ctcaacagcttgatctctacacacacacacacacacacacacacacacac 522

Db 2526 ctccac 2585

Qy 523 aactgaagctctgctctctctctctctctctctctctctctctctctct 582

Db 2586 aatcctac 2645

Qy 583 agagac 2705

Db 2646 tcatctctgagacacacacacacacacacacacacacacacacacac 2765

Qy 643 ggaac 2720

Db 2706 agttccatctctctca 2720

Query Match

Best Local Similarity 7.1%; Score 75; DB 1; Length 5940;

Matches 165; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Db 343 tcaaacctctctgacacatctgagcttgagcttgagacattacagatctca 402

Db 2406 tccgacatctgctgggagatctgctgctgggagacatctgctgctgctg 2465

Qy 403 ggcctctcaaaaacacatccaaggaactacccctgagcttgagaaatcga 462

Db 2466 gacctcagcagacacacacacacacacacacacacacacacacacac 2525

Qy 463 ctcaacagcttgatctctacacacacacacacacacacacacacacacac 522

Db 2526 ctccac 2585

Qy 523 aactgaagctctgctctctctctctctctctctctctctctctctctct 582

Db 2586 aatcctac 2645

Qy 583 agagac 2705

Db 2646 tcatctctgagacacacacacacacacacacacacacacacacacac 2765

FT	Intron	/number= 4 4430. .4528
FT		/*tag- f
FT	Intron	/number= 5 4642. .4757
FT		/*tag- g
FT	Intron	/number= 6 4890. .4967
FT		/*tag- h
FT	Intron	/number= 7 5295. .5803
FT		/*tag- i
FT	Intron	/number= 8 6197. .6339
FT		/*tag- j
FT		/number= 9
PN		W09743427-A1.
PD	20-NOV-1997.	
PF	13-MAY-1997;	E02443.
PR	14-MAY-1996; GB-010044.	
PA	(NOVS) NOVARTIS AG.	
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ; WP1: 98-086529/08.	
DR	P-BSD8; W47017.	
PT	Production of apomictic seeds - useful in plant breeding	
PS	Claim 26; Pages 64-67; 123pp; English.	
CC	The sequence is that encoding SERK, a putative receptor kinase.	
CC	It may be used as part of a method of producing apomictic seeds	
CC	comprising: (a) transforming plant material with a nucleotide	
CC	sequence encoding a protein which in active form in a cell or	
CC	cell membrane renders the cell embryogenic; (b) regenerating	
CC	the transformed material into plants or carpel-containing	
CC	plant parts; and (c) expressing the sequence in the vicinity	
CC	of the embryo sac. The apomictic seeds and embryos thus produced	
CC	can be developed into plant progeny. This is useful in plant	
CC	breeding programs. Controllable and reproducible apomixis provides	
CC	many advantages in plant improvement and cultivar development in	
CC	the case that sexual plants are available as crosses with the	
CC	apomictic plant. Apomixis provides for true-breeding, seed	
CC	propagated hybrids and could shorten and simplify the breeding	
CC	process so that selfing and progeny testing to produce and/or	
CC	stabilise a desirable gene combination could be eliminated.	
CC	Apomixis allows plant breeders to develop cultivars with	
CC	specific stable traits for such characteristics as height,	
CC	seed and forage quality and maturity.	
SQ	Sequence 4081 BP; 1120 A; 770 C; 785 G; 1406 T;	

Query Match		6.7%;	Score 70.8;	DB 1;	Length 4081;
Best Local Similarity		69.6%;	Pred No. 3.2e-10;		
Matches	96; Conservative	0;	Mismatches 42;	Indels	0;
Gaps					0

```

OY      195 aggaagatgctcttaacgtcccttcgcccgaggattaaagaatccggaccagtgttccagag 254
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1794 AGGTAGTGCTTGGCATACATTTGAGGGTACTCTAATTGCACAACAAATGTCTTGCAGAG 1853
OY      255 cggagatccaactctgtttaatcttgttaaccttggtcccatgtaactgttaacaagaacaa 314
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1854 CTGGGATCTTACCCTTGATGAATCCTTGACACATGGTTCACATGTCTTGCAACAACGAGA 1913
OY      315 ccgcgcaactcactgtgtga 332
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1914 CAGTGTCAATAGAGTGTGA 1931

```

RESULT	13
ID	V06570
AC	V06570 standard; DNA; 6695 BP.
DT	03-ANG-1998 (first entry)
DE	Daucus carota SERK gene.
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos; plant breeding; ds.

OS	Daucus carota.	Location/Qualifiers
PH	Key	3696..6620
FT	CDS	/tag= a /note= "contains introns"
FT		3731..3802
FT	intron	/tag= b /number= 1
FT		3851..3979
FT	intron	/tag= c /number= 2
FT		4124..4211
FT	intron	/tag= d /number= 3
FT		4284..4357
FT	intron	/tag= e /number= 4
FT		4430..4528
FT	intron	/tag= f /number= 5
FT		4642..4757
FT	intron	/tag= g /number= 6
FT		4890..4967
FT	intron	/tag= h /number= 7
FT		5295..5803
FT	intron	/tag= i /number= 8
FT		6197..6339
FT	intron	/tag= j /number= 9
FT		
PN	WO9743427-A1.	
PD	20-NOV-1997.	
PF	13-MAY-1997; E02443.	
PR	14-MAY-1996; GB-010044.	
PA	(NOVS) NOVARTIS AG.	
P1	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ.	
DR	WP1: 98-086529/08.	
P-PDB:	W47013..	
PS	Production of apomictic seeds - useful in plant breeding	
CS	Claim 21: Pages 40-46; 123pp; English.	
CC	The sequence is that encoding SERK, a putative receptor kinase.	
CC	It may be used as part of a method of producing apomictic seeds.	
CC	comprising: (a) transforming plant material with a nucleotide	
CC	sequence encoding a protein which in active form in a cell or	
CC	cell membrane renders the cell embryogenic; (b) regenerating	
CC	the transformed material into plants or carpel-containing	
CC	plant parts; and (c) expressing the sequence in the vicinity	
CC	of the embryo sac. The apomictic seeds and embryos thus produced	
CC	can be developed into plant progeny. This is useful in plant	
CC	breeding programs. Controllable and reproducible apomixis provides	
CC	many advantages in plant improvement and cultivar development in	
CC	the case that sexual plants are available as crosses with the	
CC	apomictic plant. Apomixis provides for true-breeding, seed	
CC	propagated hybrids and could shorten and simplify the breeding	
CC	process so that selfing and progeny testing to produce and/or	
CC	stabilise a desirable gene combination could be eliminated.	
CC	Apomixis allows plant breeders to develop cultivars with	
CC	specific stable traits for such characteristics as height,	
CC	seed and forage quality and maturity.	
SQ	Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;	

Query Match 6.7%; Score 70.8; DB 1; Length 6695;
 Best Local Similarity 69.6%; Pred. No. 3.9e-10;
 Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY	195 aggaagtgcctttagcccttcgacggcggaatacatcgacatgtttctccagag	254
DB	956 AGCGAGTGCCTTTACAAACTTAGCAACTGACGTTCGGCAAGATCCCAACAATGTCTCTCAGAG	1015
OY	255 ctggagatccaactcttgtaaatcccttgtaacctggtgtccatgtcacctgtaaccagaacaa	314

Db 1016 CTGGATCCACACCTGTGACATGCTTCTGACATGTACAAATGAAA 1075
QY 315 ccgcgcactcgtgtgga 332
Db 1076 CAGTGTATAGAGTSTA 1093

RESULT 14

VL4518
ID VL4518 standard; DNA; 3979 BP.
AC VL4518;
DT 20-MAY-1998 (first entry)
DE CF-5 pathogen resistance gene variant #1.
KW Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon pimpinellifolium.
FH Key Location/Qualifiers
FT CDS 653..3560
FT /tag= a
FT /product= CF-5 pathogen resistance gene
PN MO9743429-A1.
PD 20-NOV-1997.
PE 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1996; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR WPI; 98-008895/01.
DR P-PSDB; W41309.
PT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT for production of transgenic plants resistant to pathogens e.g.
PT tomato leaf mould C. fulvum in tomatoes
PS Claim 3: Fig 1a: 75bp: English.
CC This sequence is an example of the polynucleotide of the invention, and
CC is able to confer pathogen resistance on a plant. It is one of two tomato
CC Cf-5 gene variants, which offer resistance against the pathogen
CC Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC the gene into plant cells and regenerating plants from the cells;
CC asexually or sexually produced offspring can also be subsequently
CC produced. Expression of the gene in plant cells can confer pathogen
CC resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC Oligonucleotides with sequences complementary to the gene or fragments
CC of it, are useful in anti-sense techniques to reduce gene expression. The
CC nucleic acids/polynucleotides are useful as hybridisation probes to
CC identify other genes/fragments conferring pathogen resistance on plants
CC e.g. Phytophthora resistance in potatoes. Homologies between Cf-5 and
CC Cf-9 may be used to identify further resistance genes of this class.
SQ Sequence 3979 BP; 1217 A; 717 C; 698 G; 1347 T;

Query Match 6.6%; Score 70.6; DB 1; Length 3979;
Best Local Similarity 50.4%; Pred. No. 3.6e-10;
Matches 172; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 313 aaccgcgcactcgtgtggaattgggaattcaacctctcgacaattcgcgcctag 372
Db 1944 AACAACTTGTATTGTTGATCTTTACAAATATACAGCTTTCGGCTCATTCCTAAGAA 2003
QY 373 ctgggaagctgaacattacagatctagagcctacaacaaacatccaagaact 432
Db 2004 ATAGGTTACCTGAGTCTCTTACTGAACTATTTTGGTAATATCTCTTATATGCTCT 2063
QY 433 ataccctccgaacttggaaatctcagaatctcacaagctgtacacaacaat 492
Db 2064 ATTCTGCTTCATTTGGGAACTAAACAACCTGTGAGGTGATACATTACAAATATAG 2123
QY 493 cttaacggagtagttccacattcttgggaataatggaatctcgtcttttaagcctt 552
Db 2124 CTTTCTGGCTCATTTCTGTTGCAATATGAAATCTCGAAACTCTGTTTCTC 2183
QY 553 aatgaacacgcatgacgggccaatccctagagcactacgcgaatcccaagccttaa 612

Db 2184 AGTGATACGATCTCATTTGGGAATTCCTTATTTGTGTGCAATTTGACATCTGGA 2243
QY 613 gtgttgatgctcgaagcaatgttgtggaacaatccc 653
Db 2244 GTGTTGATATGTGCGAAGAAACATTTGAAGGAAAGTTCC 2284

RESULT 15

VL4519
ID VL4519 standard; DNA; 3979 BP.
AC VL4519;
DT 20-MAY-1998 (first entry)
DE CF-5 pathogen resistance gene variant #2.
KW Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon pimpinellifolium.
FH Key Location/Qualifiers
FT CDS 653..3560
FT /tag= a
FT /product= CF-5 pathogen resistance gene
PN MO9743429-A1.
PD 20-NOV-1997.
PE 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1996; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR WPI; 98-008895/01.
DR P-PSDB; W41310.
PT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT for production of transgenic plants resistant to pathogens e.g.
PT tomato leaf mould C. fulvum in tomatoes
PS Claim 6: Fig 1b: 75bp: English.
CC This sequence is an example of the polynucleotide of the invention, and
CC is able to confer pathogen resistance on a plant. It is one of two tomato
CC Cf-5 gene variants, which offer resistance against the pathogen
CC Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC the gene into plant cells and regenerating plants from the cells;
CC asexually or sexually produced offspring can also be subsequently
CC produced. Expression of the gene in plant cells can confer pathogen
CC resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC Oligonucleotides with sequences complementary to the gene or fragments
CC of it, are useful in anti-sense techniques to reduce gene expression. The
CC nucleic acids/polynucleotides are useful as hybridisation probes to
CC identify other genes/fragments conferring pathogen resistance on plants
CC e.g. Phytophthora resistance in potatoes. Homologies between Cf-5 and
CC Cf-9 may be used to identify further resistance genes of this class.
SQ Sequence 3979 BP; 1218 A; 716 C; 698 G; 1347 T;

Query Match 6.6%; Score 70.6; DB 1; Length 3979;
Best Local Similarity 50.4%; Pred. No. 3.6e-10;
Matches 172; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 313 aaccgcgcactcgtgtggaattgggaattcaacctctcgacaattcgcgcctag 372
Db 1944 AACAACTTGTATTGTTGATCTTTACAAATATACAGCTTTCGGCTCATTCCTAAGAA 2003
QY 373 ctgggaagctgaacattacagatctagagcctacaacaaacatccaagaact 432
Db 2004 ATAGGTTACCTGAGTCTCTTACTGAACTATTTTGGTAATATCTCTTATATGCTCT 2063
QY 433 ataccctccgaacttggaaatctcagaatctcacaagctgtacacaacaat 492
Db 2064 ATTCTGCTTCATTTGGGAACTAAACAACCTGTGAGGTGATACATTACAAATATAG 2123
QY 493 cttaacggagtagttccacattcttgggaataatggaatctcgtcttttaagcctt 552
Db 2124 CTTTCTGGCTCATTTCTGTTGCAATATGAAATCTCGAAACTCTGTTTCTC 2183
QY 553 aatgaacacgcatgacgggccaatccctagagcactacgcgaatcccaagccttaa 612
Db 2184 AGTGATACGATCTCATTTGGGAAATTCCTTATTTGTGTGCAATTTGACATCTGGA 2243

Qy 613 gtgttgatgtctcaagcaatgtattgtgtggaacaatccc 653
| | | | | | | | | | | | | | | | | | | | | |
Db 2244 GTGTTGTATATGTCGAGAAACAATTGGAAGGAAAAAGTTCC 2284

Search completed: June 23, 2000, 22:53:58
Job time: 40634 sec